BlockFlex: Enabling Storage Harvesting with Software-Defined Flash in Modern Cloud Platforms

Benjamin Reidys* Jinghan Sun*

Anirudh Badam† Shadi Noghabi† Jian Huang

*Co-Primary Authors

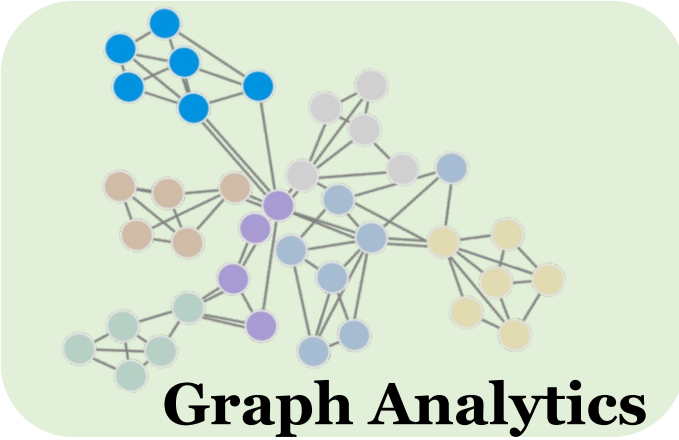
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

†

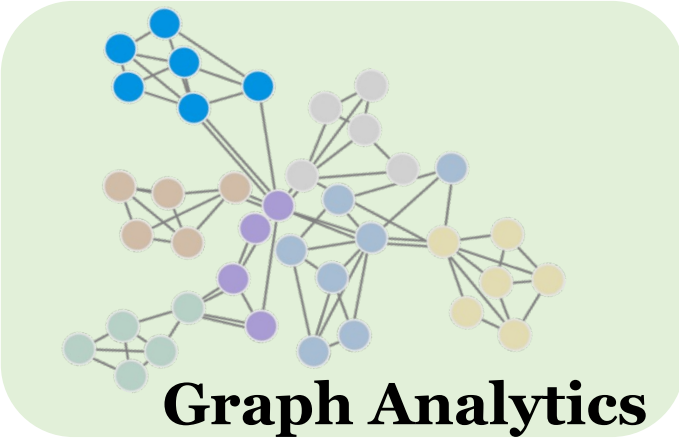
Microsoft

Cloud Storage: A Core Service in Data Centers Today

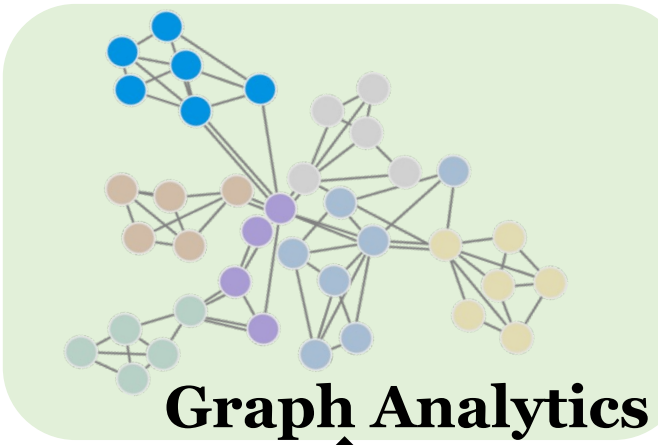
Cloud Storage: A Core Service in Data Centers Today



Cloud Storage: A Core Service in Data Centers Today

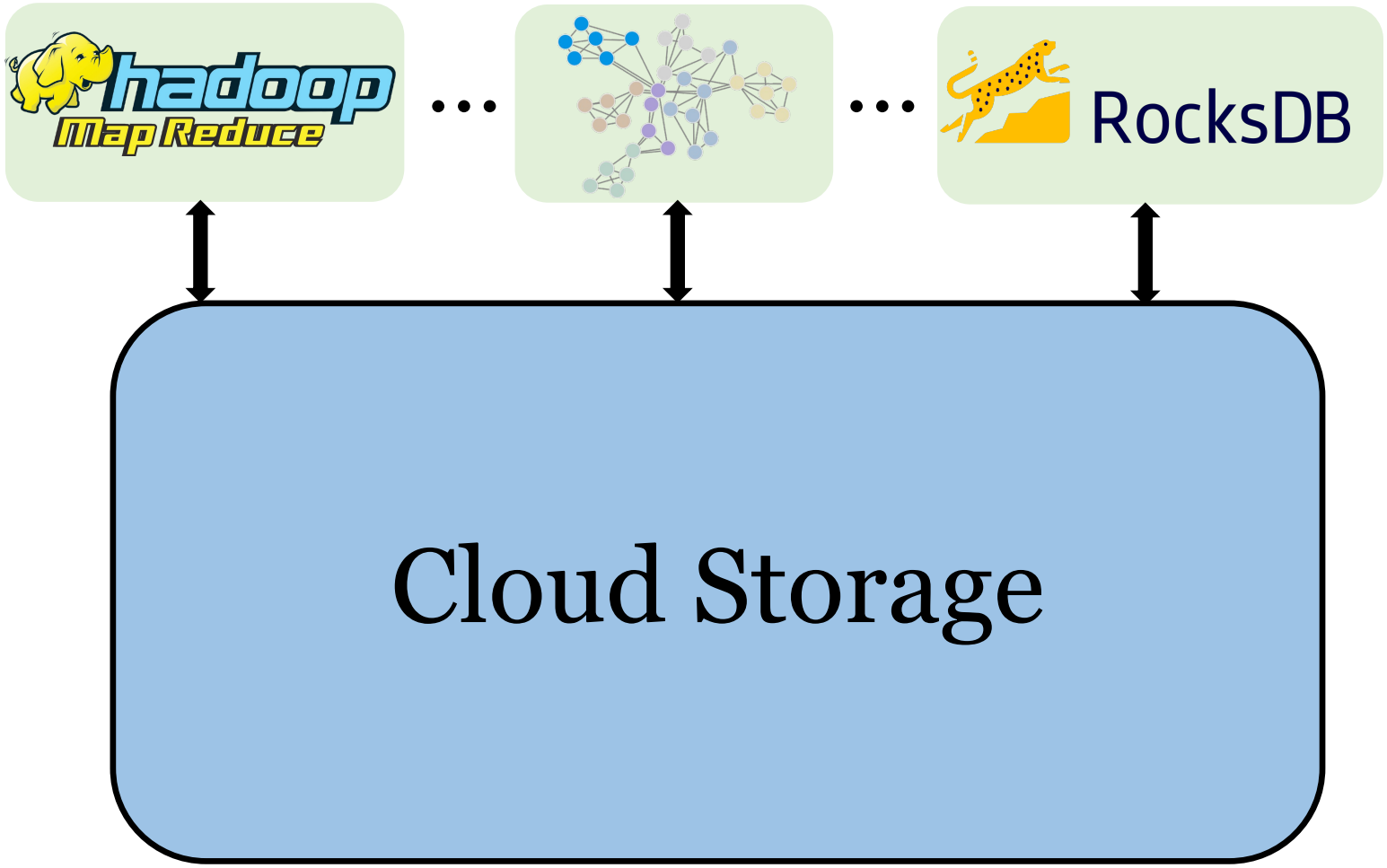


Cloud Storage: A Core Service in Data Centers Today

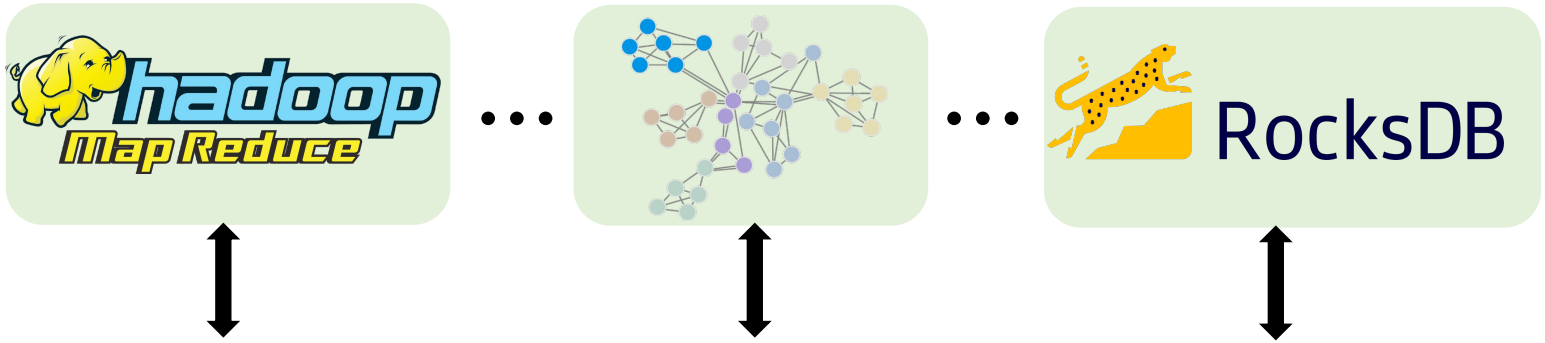


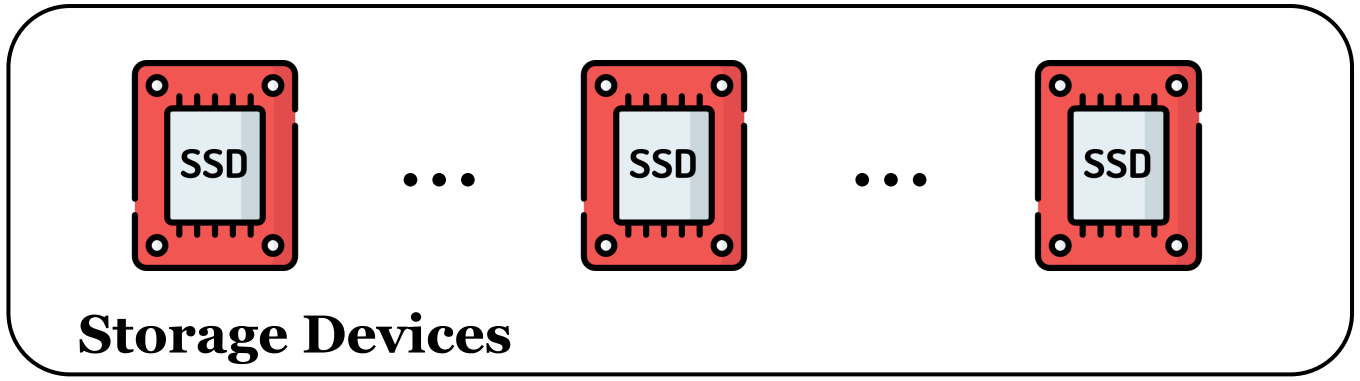
Cloud Storage

Storage Virtualization: The Backbone for Cloud Services

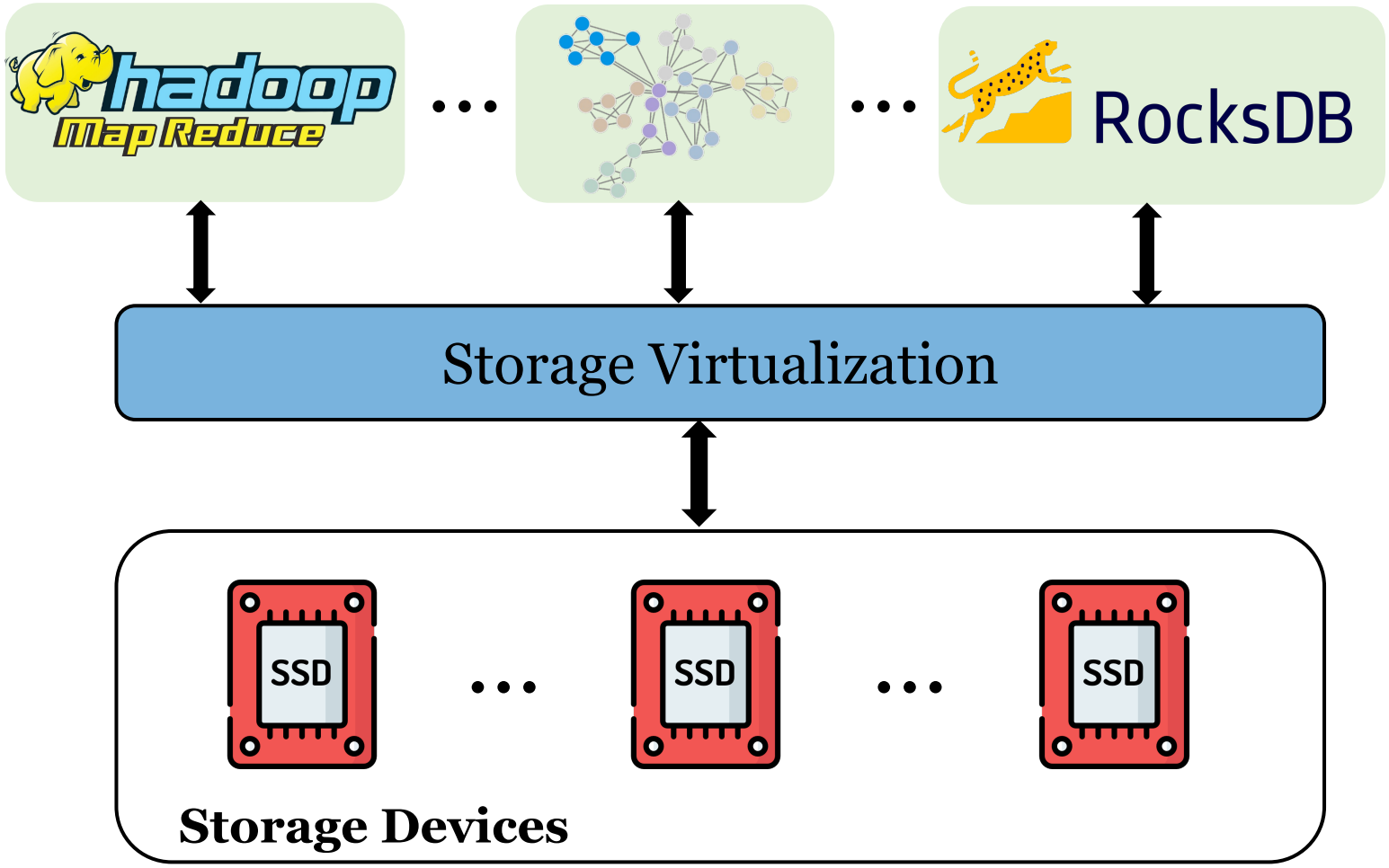


Storage Virtualization: The Backbone for Cloud Services

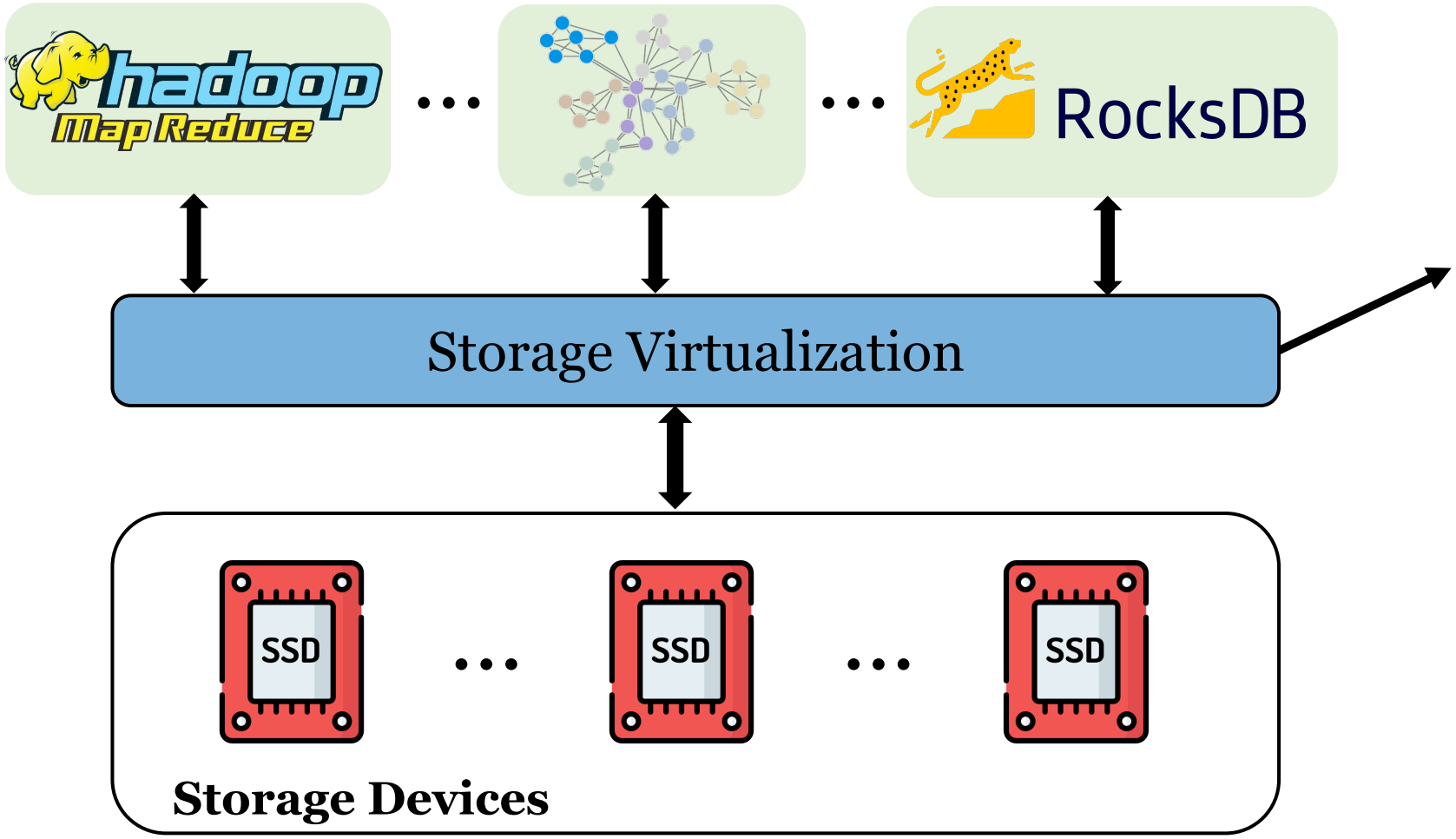


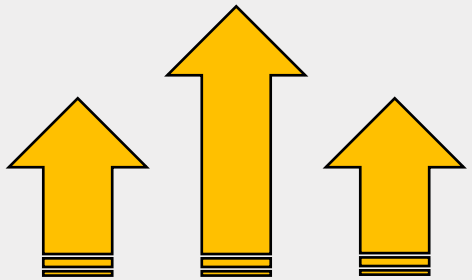


Storage Virtualization: The Backbone for Cloud Services



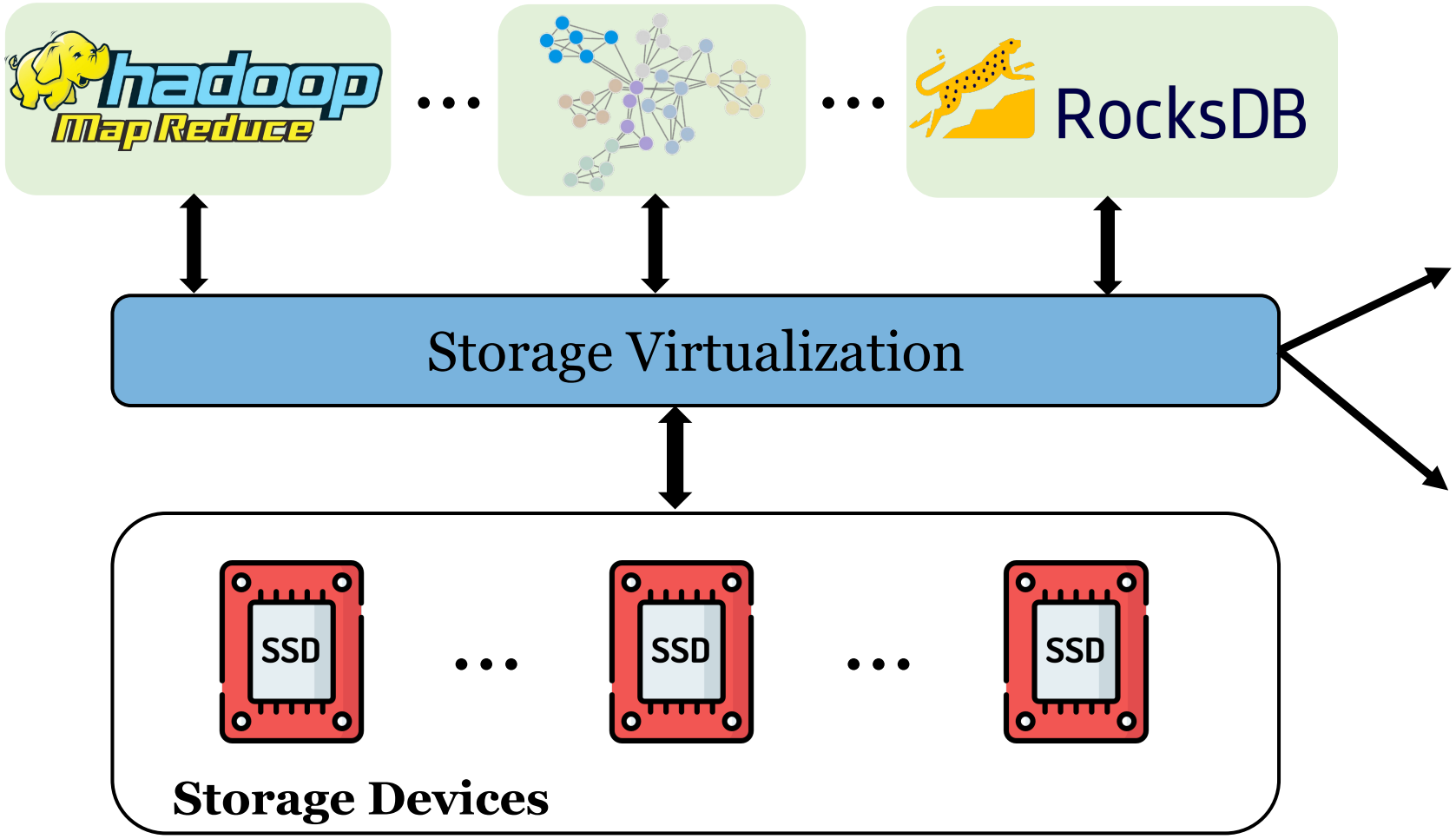
Storage Virtualization: The Backbone for Cloud Services

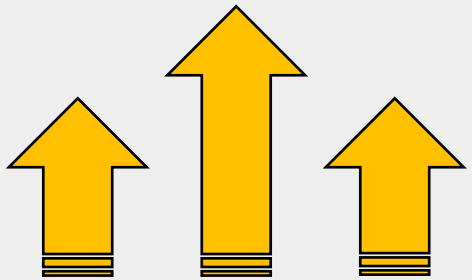




Resource Sharing for
Improved Utilization

Storage Virtualization: The Backbone for Cloud Services



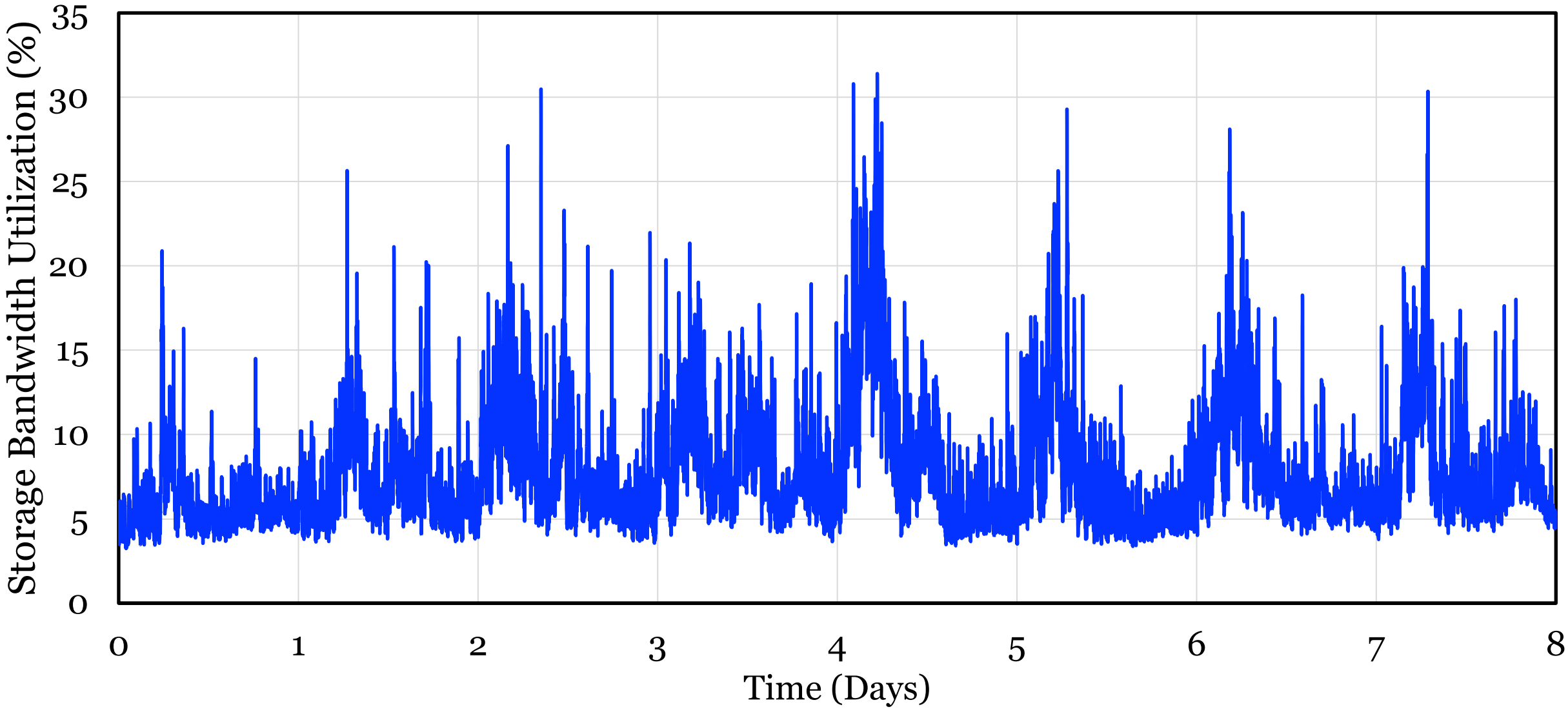


Resource Sharing for **Improved Utilization**

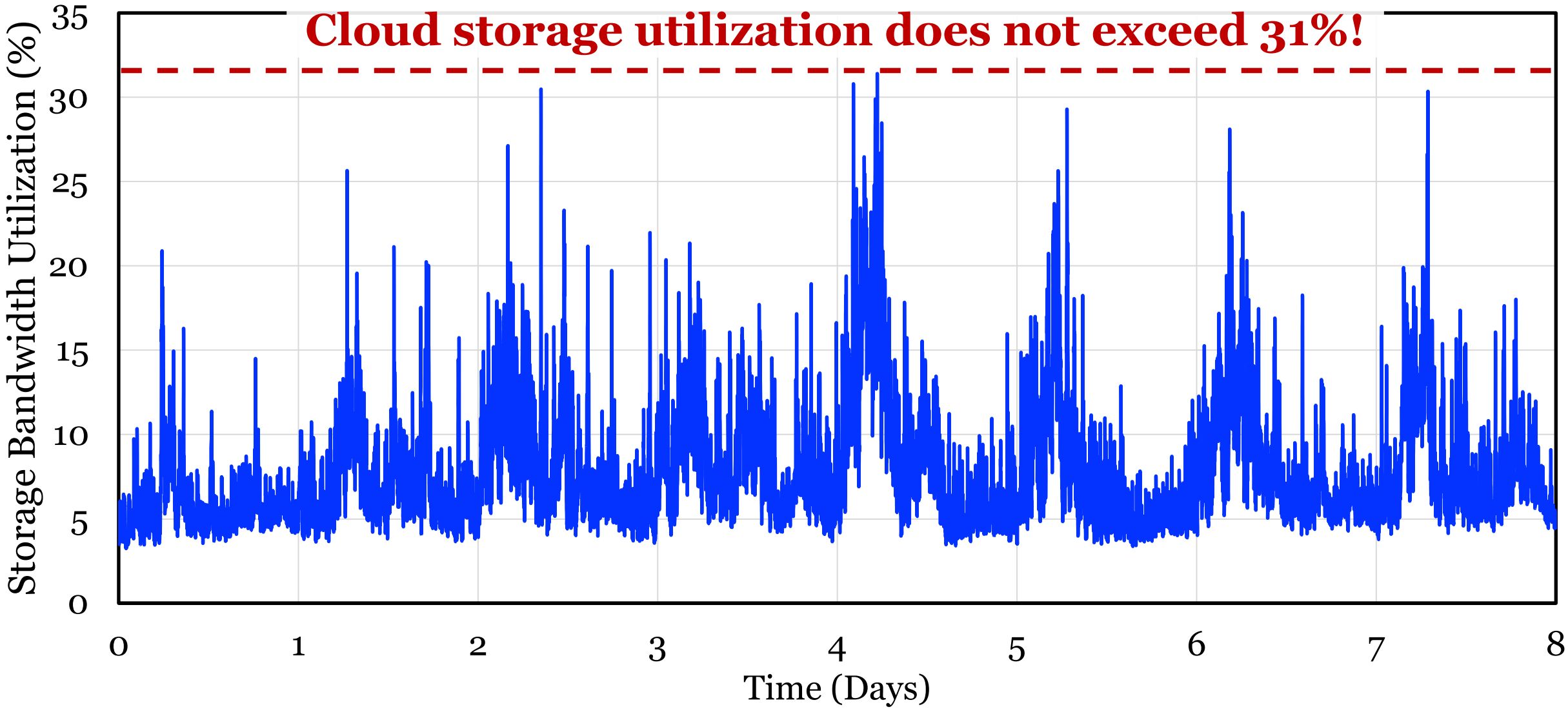
Storage Complexity is Hidden from the User

Cloud Storage is Severely Underutilized

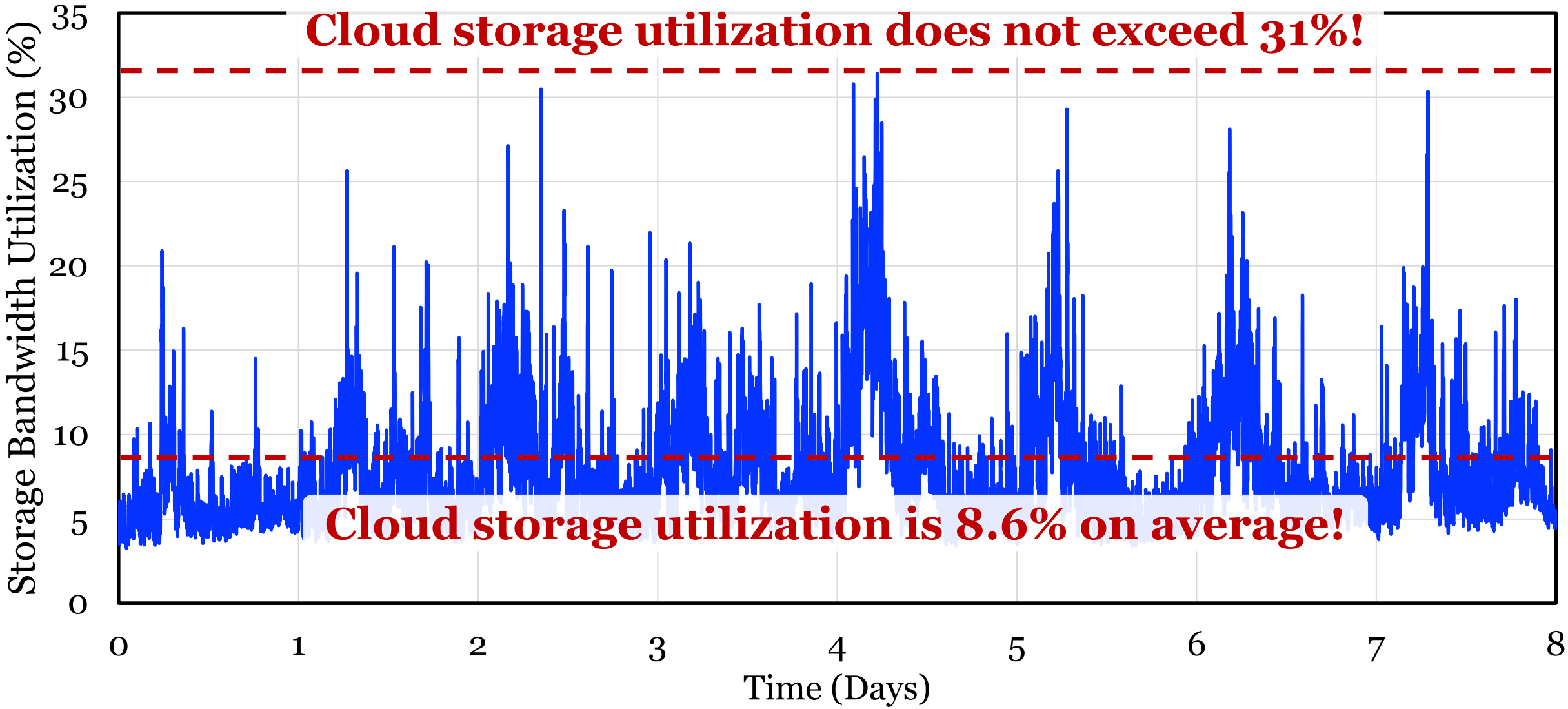
Cloud Storage is Severely Underutilized



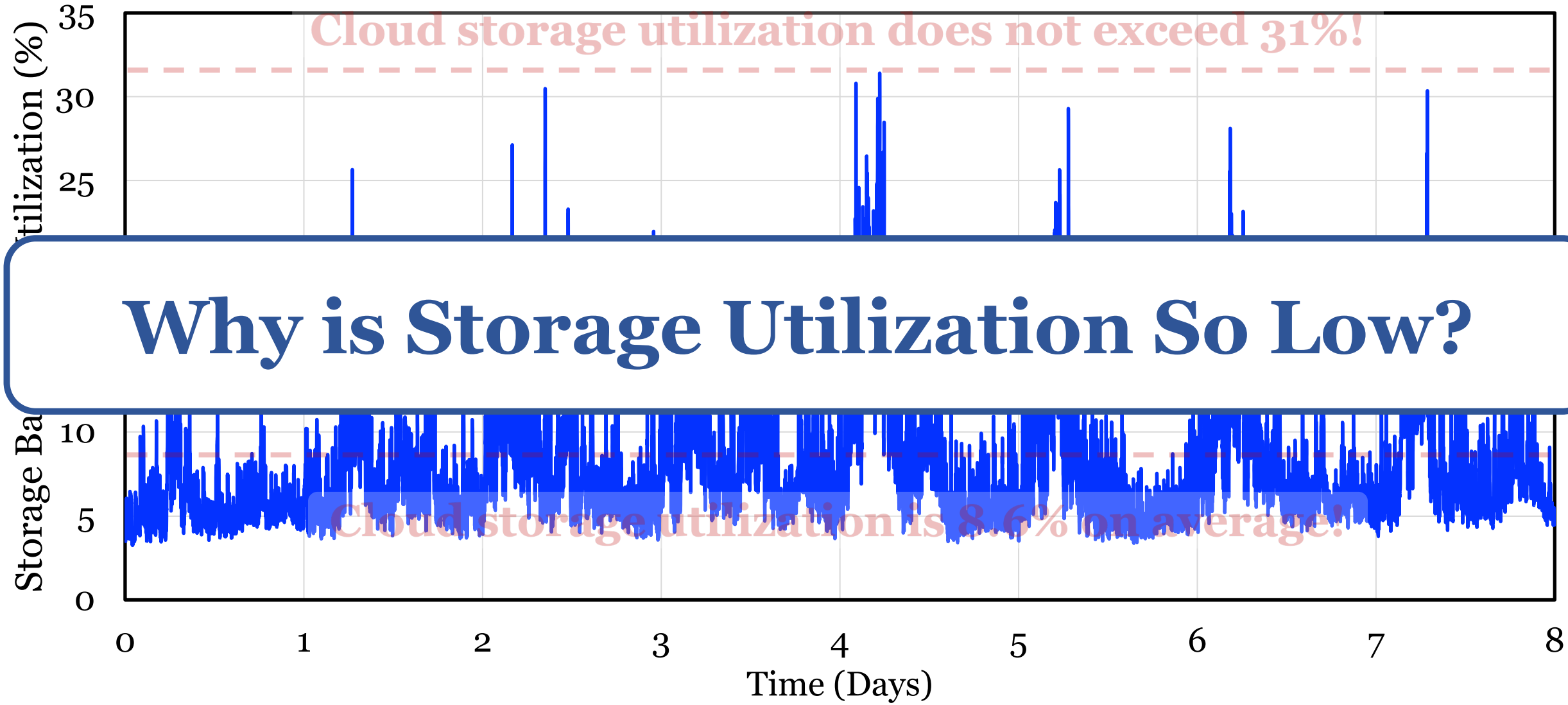
Cloud Storage is Severely Underutilized



Cloud Storage is Severely Underutilized

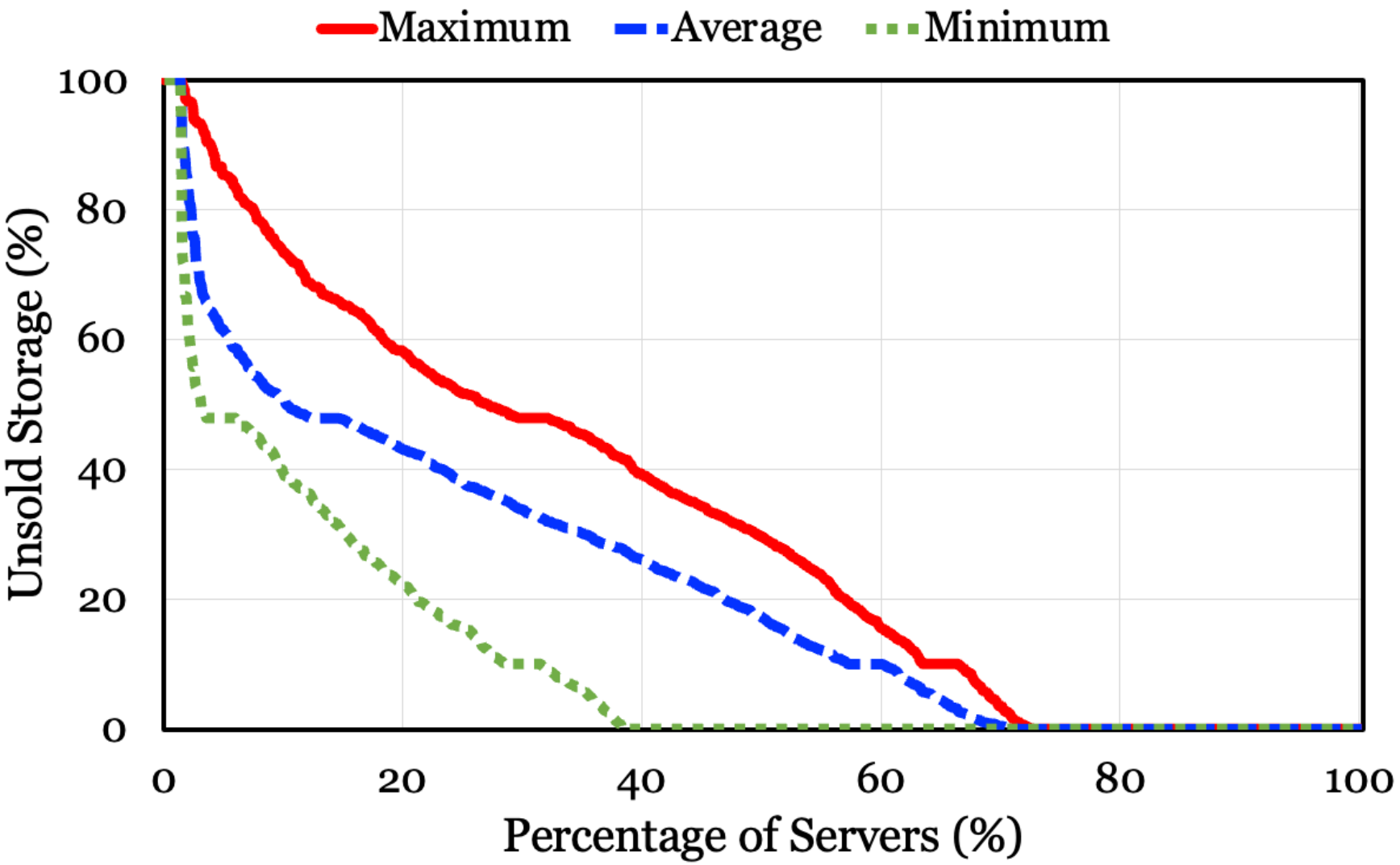


Cloud Storage is Severely Underutilized

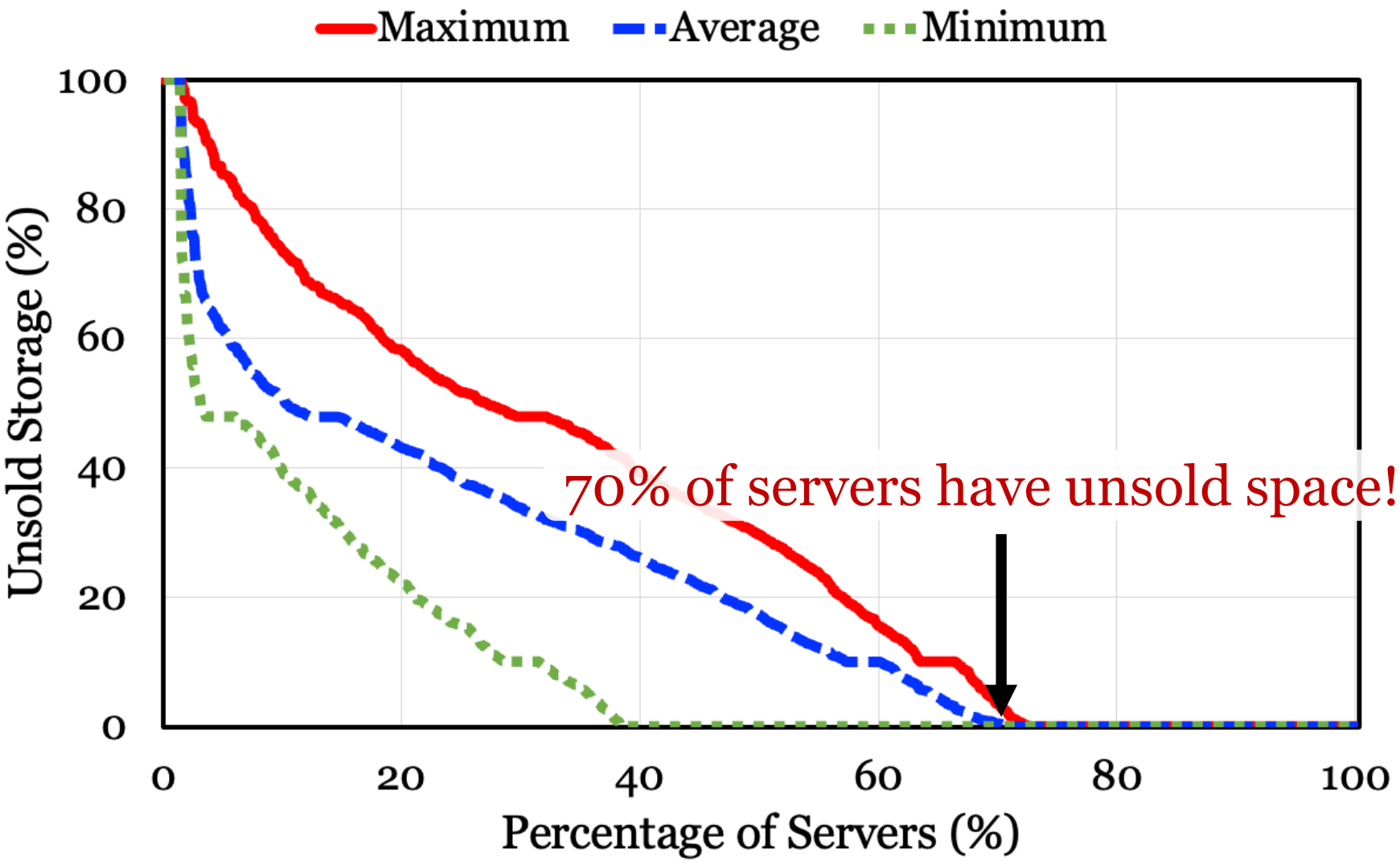


Why is Storage Utilization So Low?

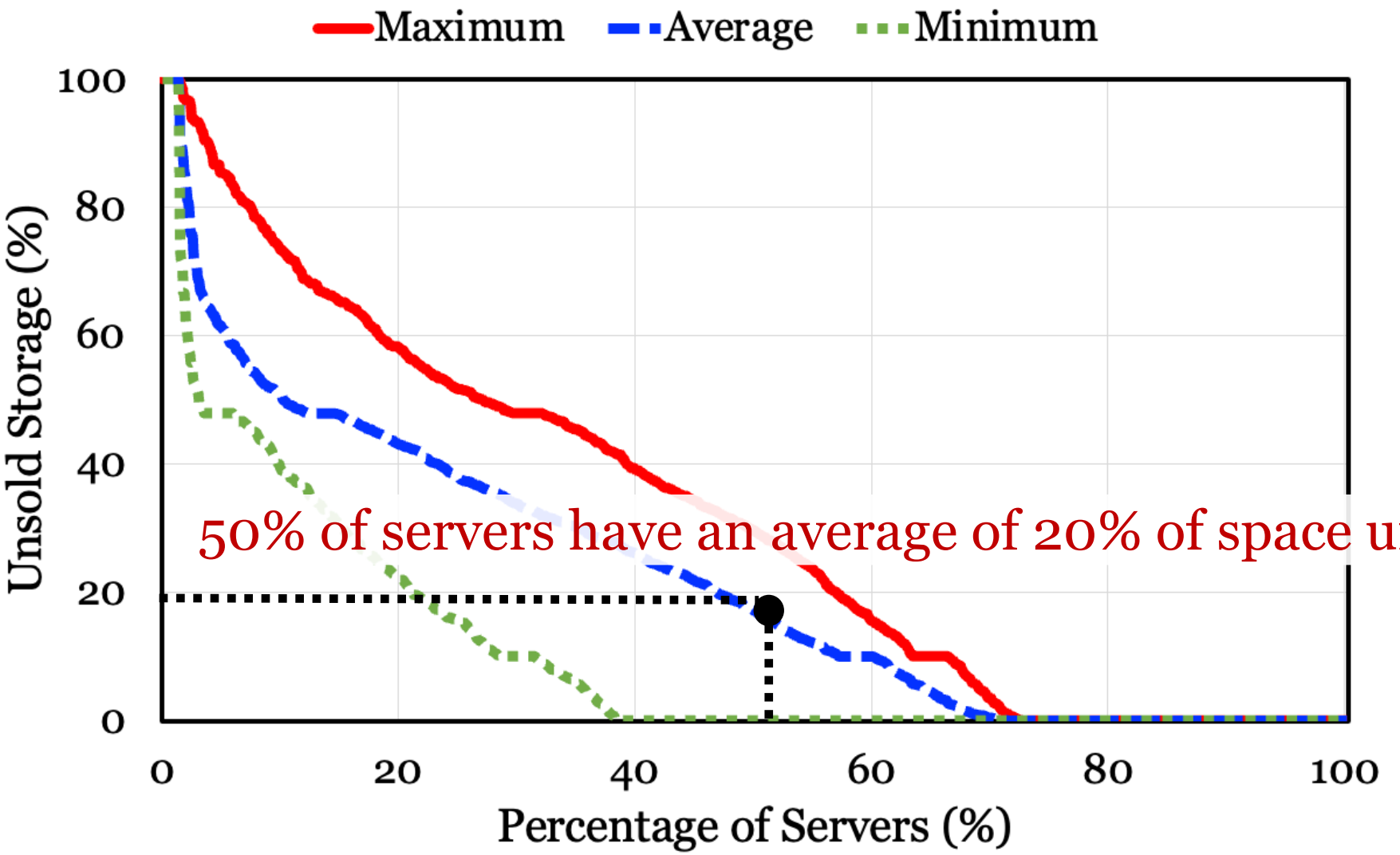
Cause 1: Low Utilization from Unallocated Storage



Cause 1: Low Utilization from Unallocated Storage

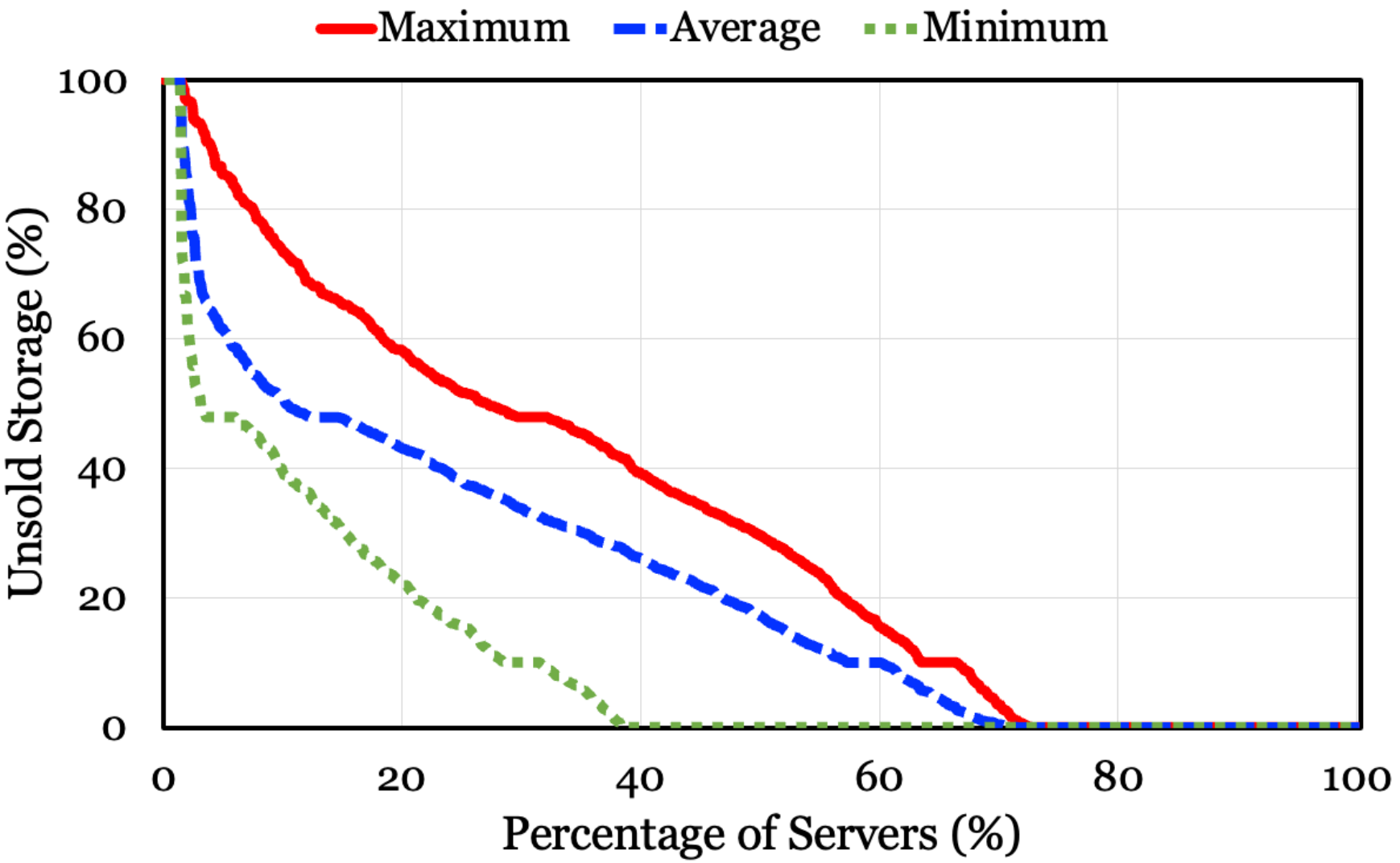


Cause 1: Low Utilization from Unallocated Storage



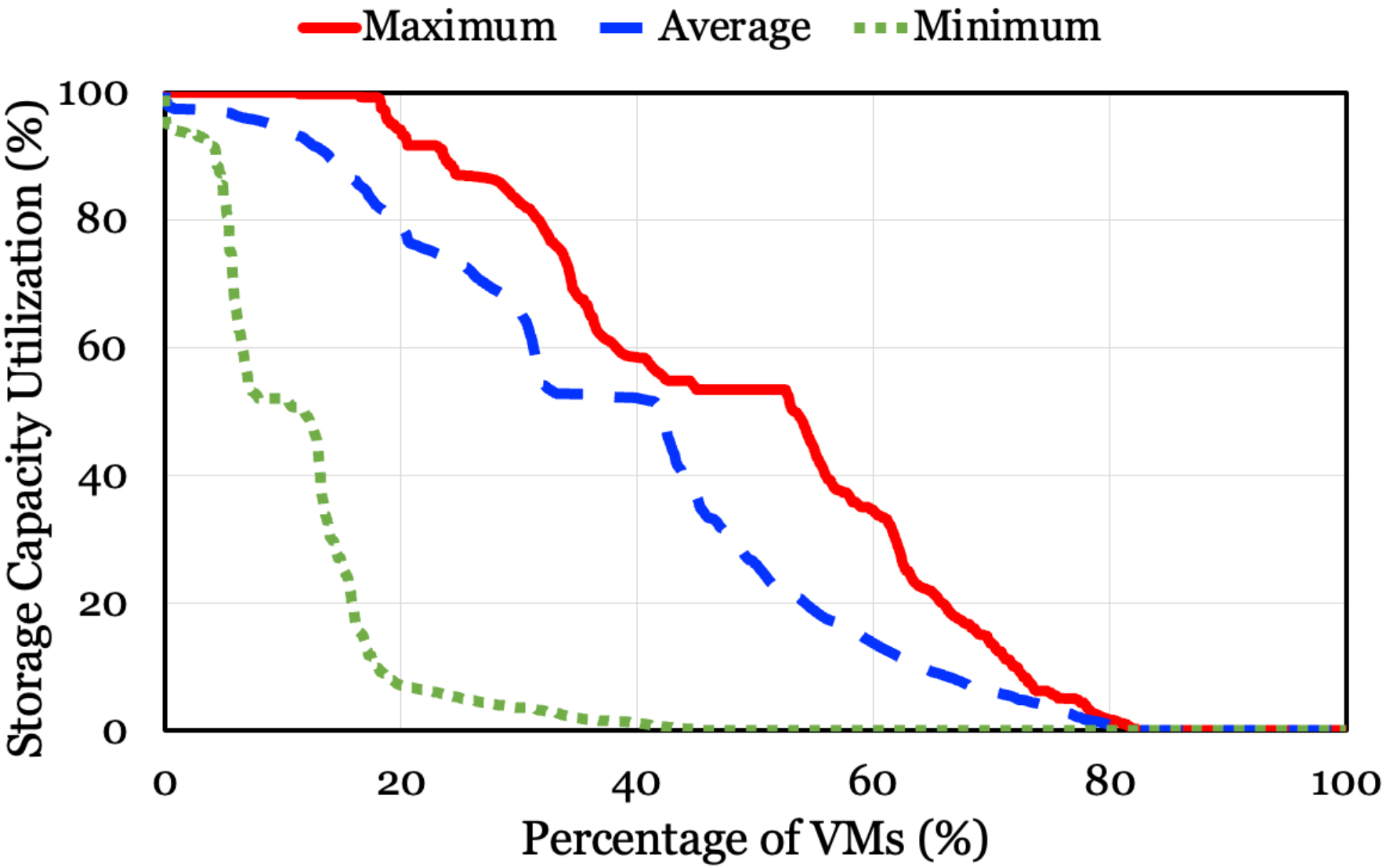
50% of servers have an average of 20% of space unsold!

Cause 1: Low Utilization from Unallocated Storage

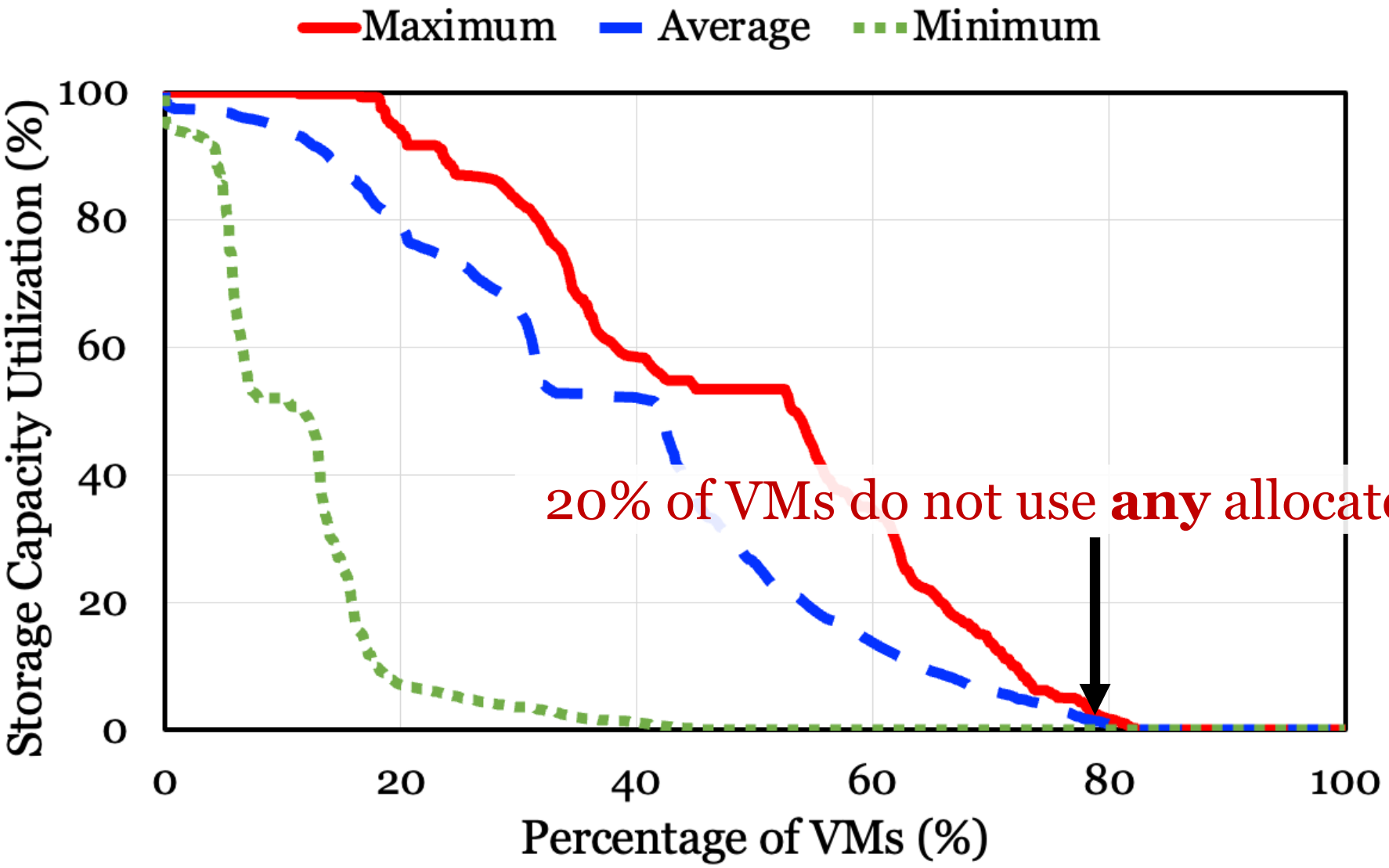


Overprovisioned Resources for VM Allocations!

Cause 2: Low Utilization from Allocated Storage

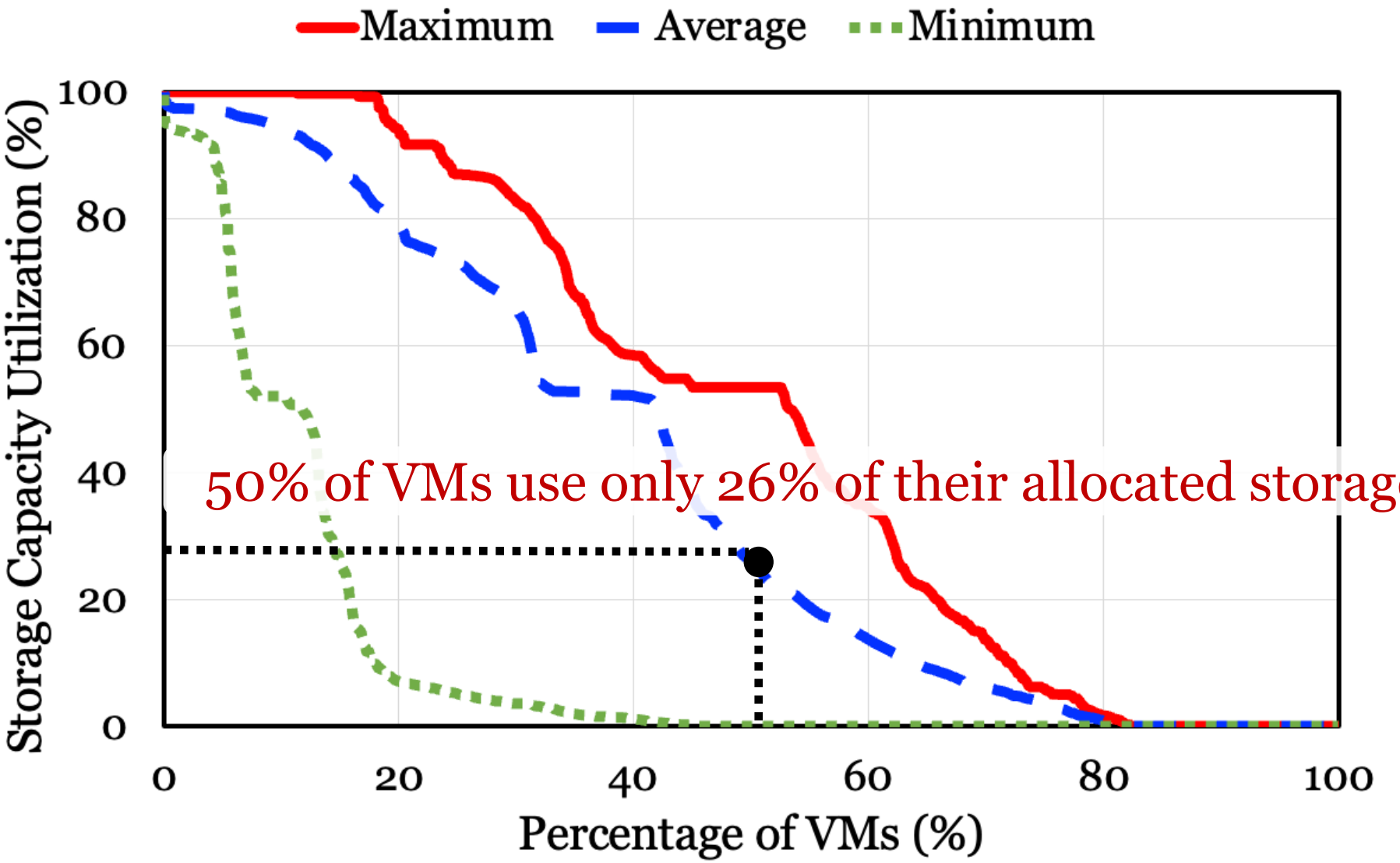


Cause 2: Low Utilization from Allocated Storage



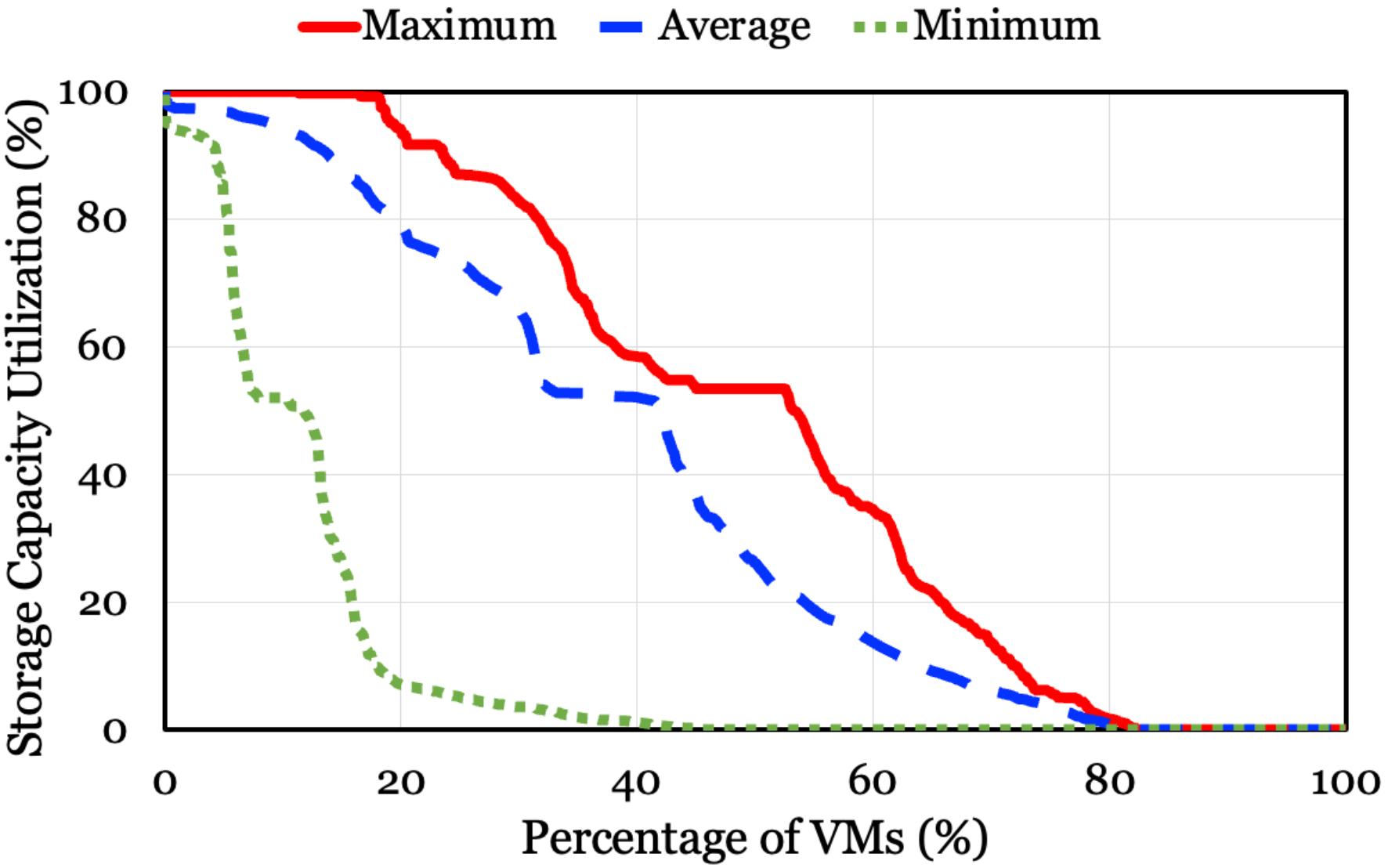
20% of VMs do not use **any** allocated storage!

Cause 2: Low Utilization from Allocated Storage



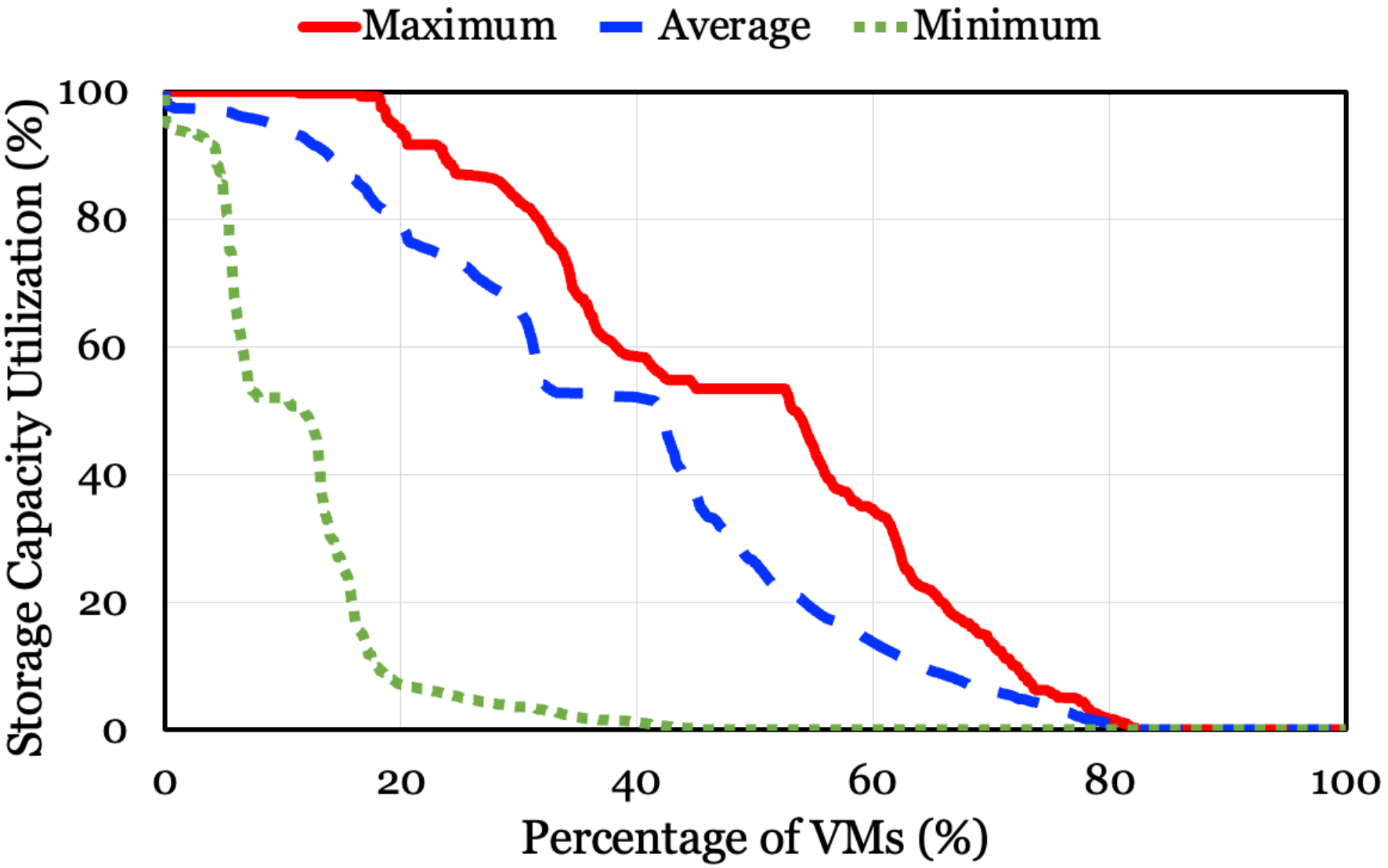
50% of VMs use only 26% of their allocated storage on average!

Cause 2: Low Utilization from Allocated Storage

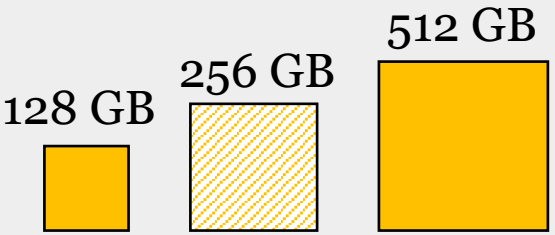


Users may **Provision for Bursts** in Workloads

Cause 2: Low Utilization from Allocated Storage

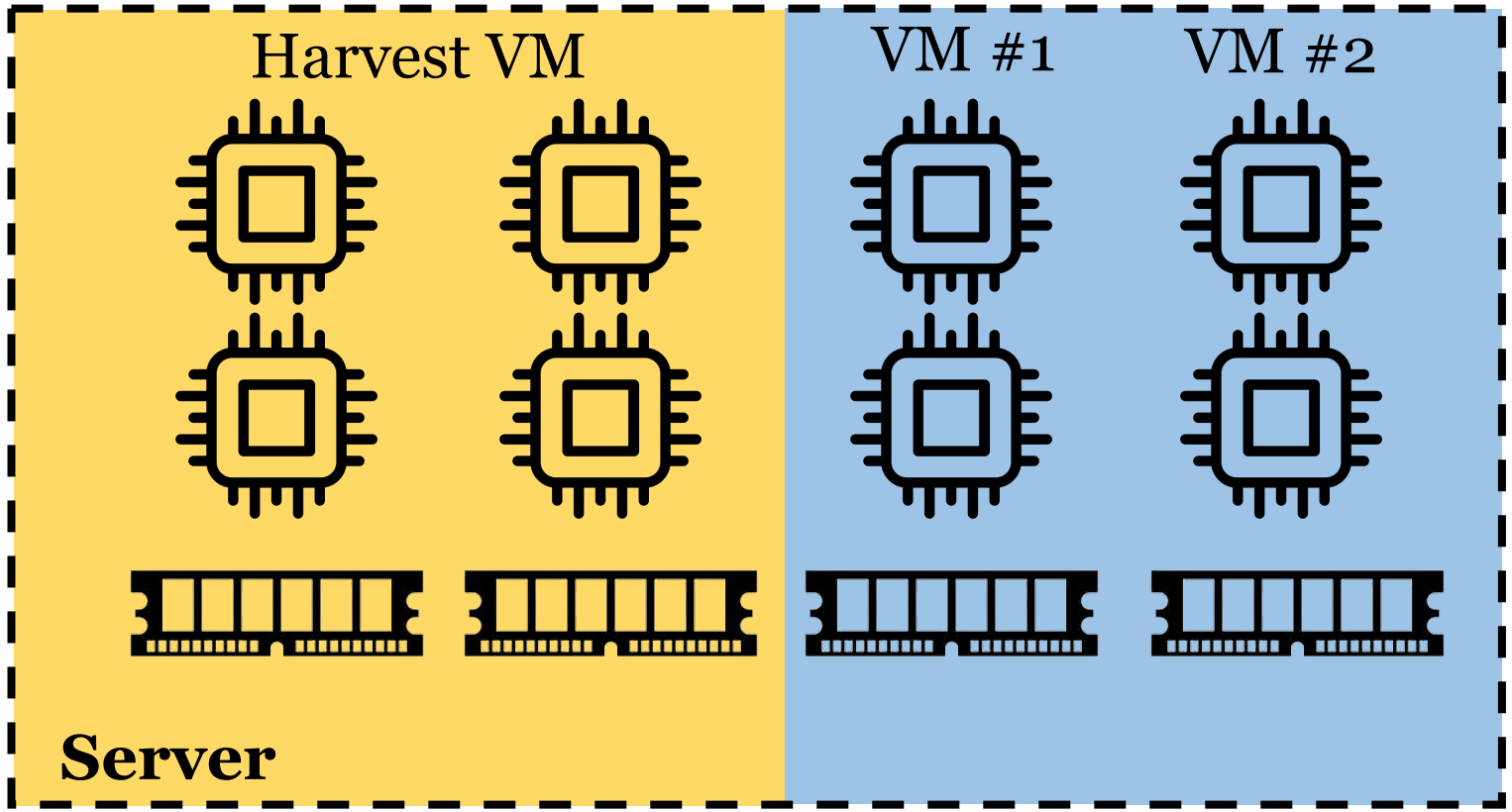


Users may **Provision for Bursts** in Workloads



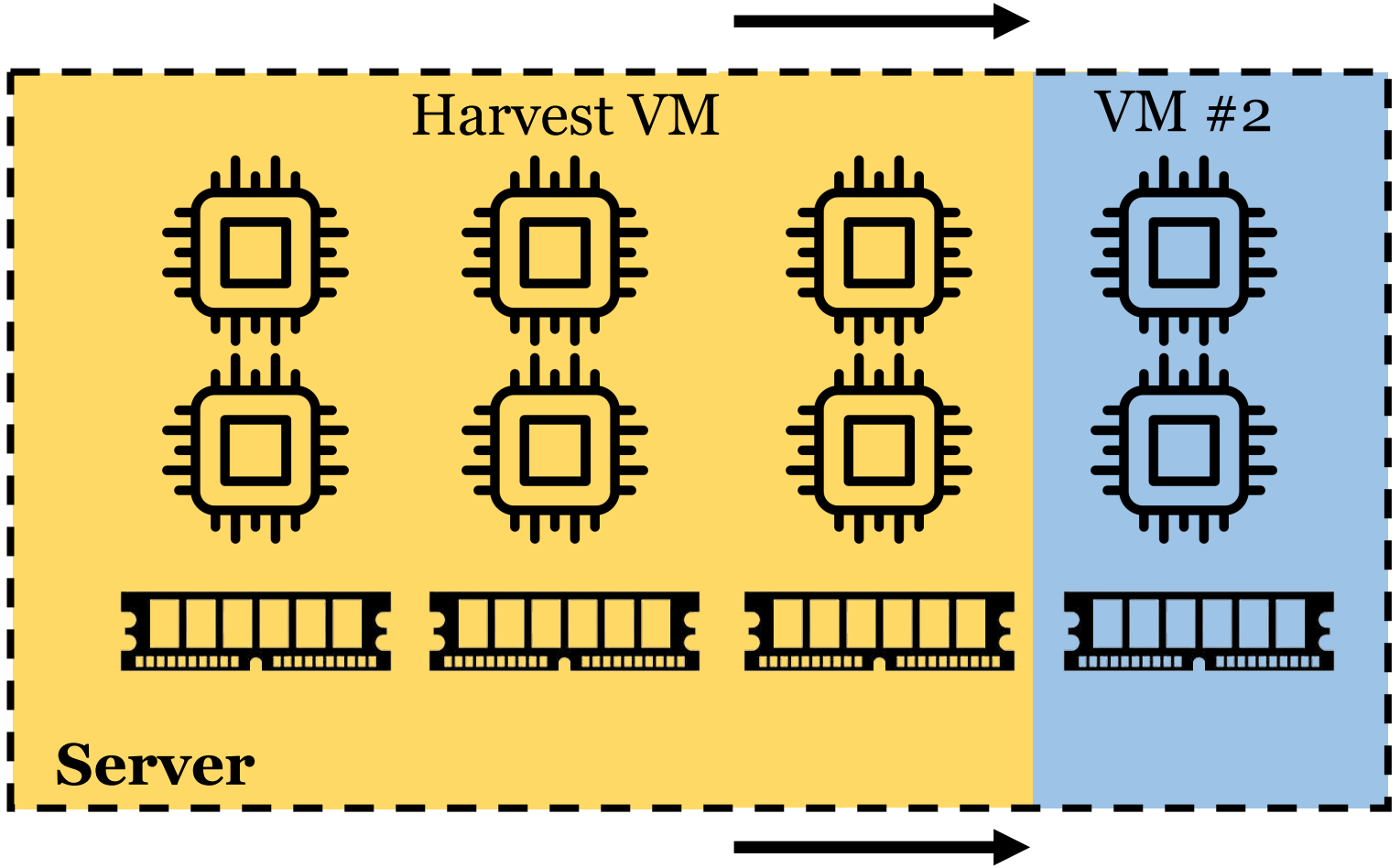
Limited Capacity Options for VM Storage Allocations

Harvest VMs: Promising for Improved Resource Utilization



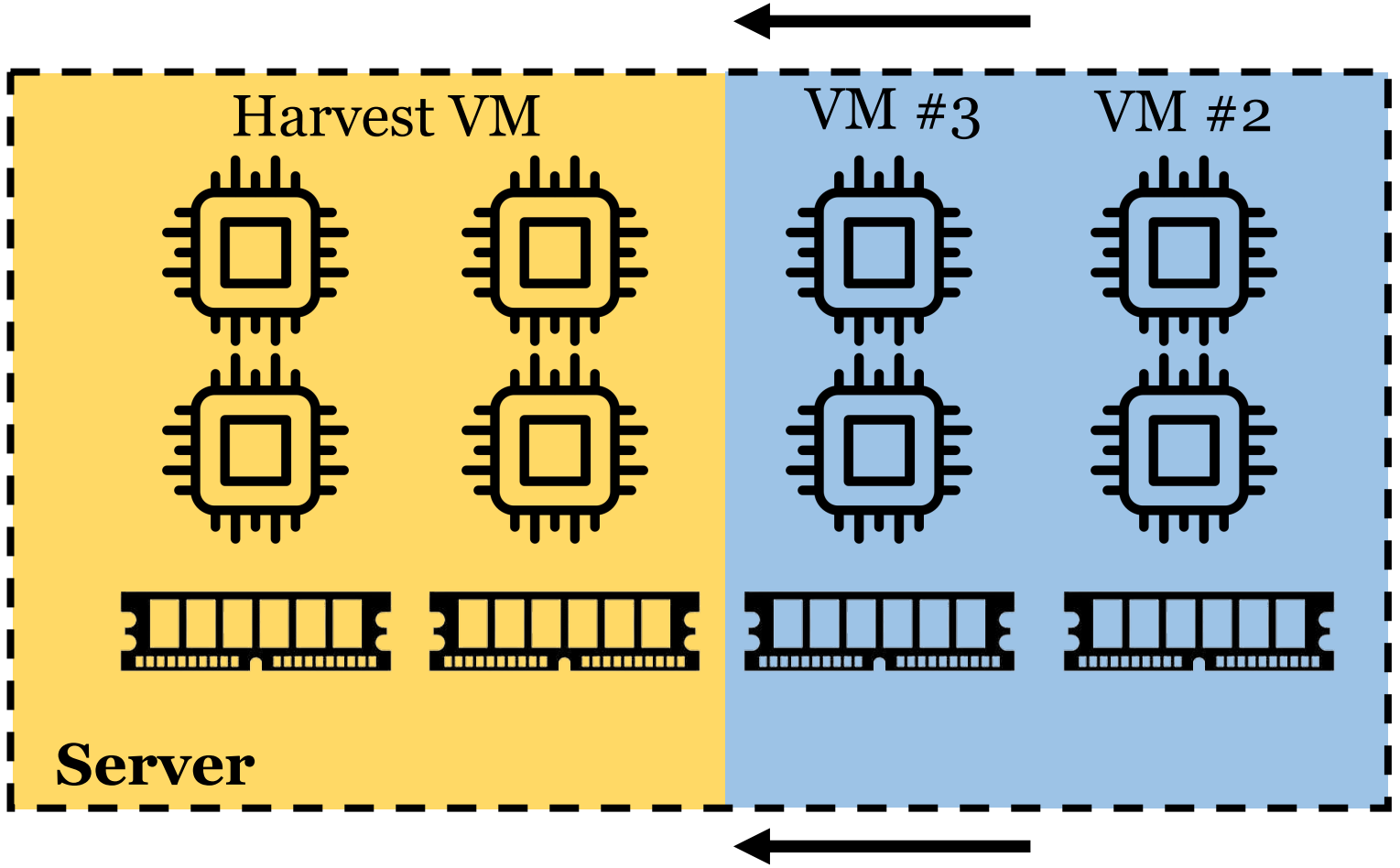
Harvest VMs: Promising for Improved Resource Utilization

Harvest VMs **grow** to **utilize** available resources!



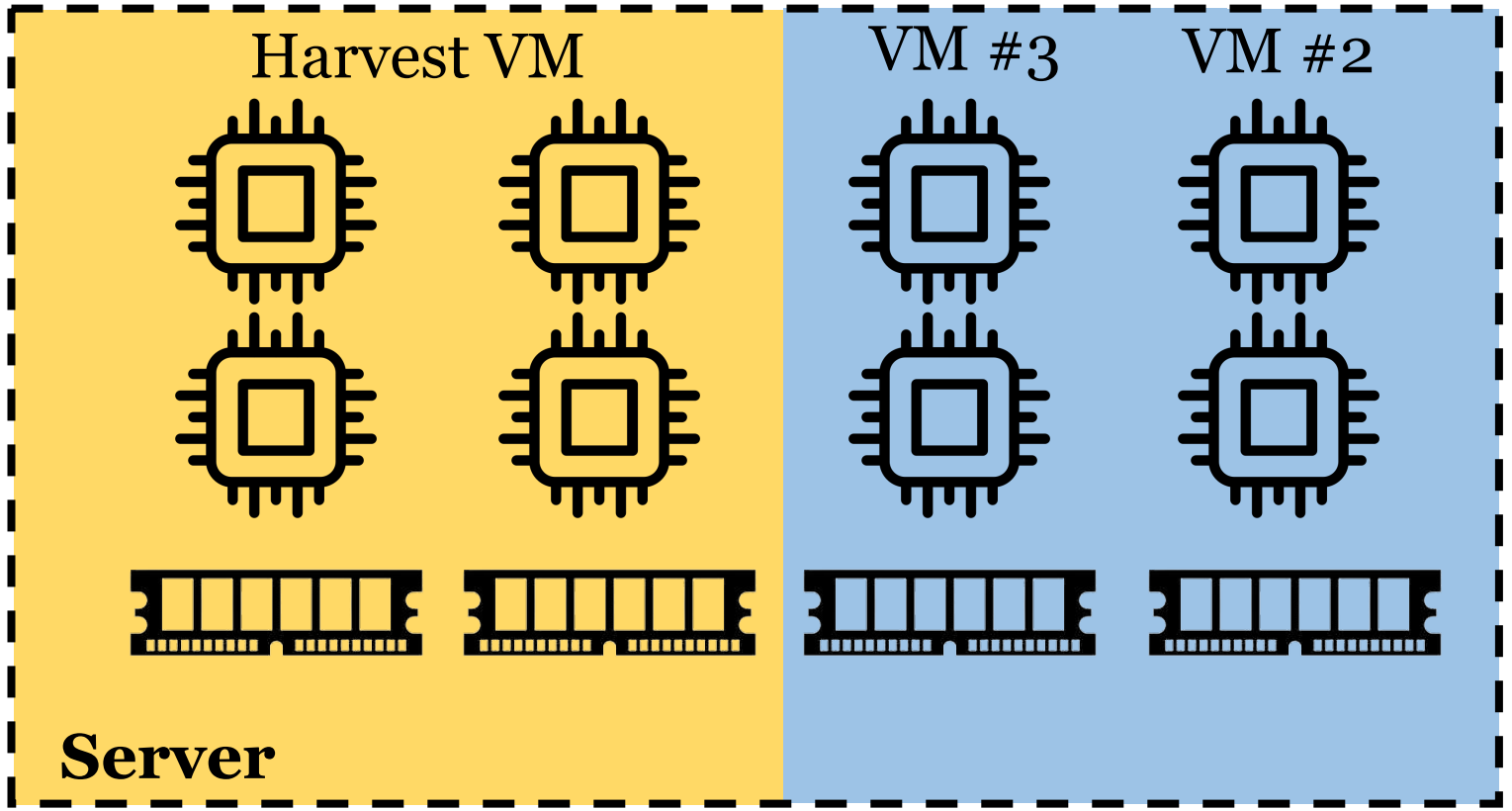
Harvest VMs: Promising for Improved Resource Utilization

Harvest VMs **shrink** to **avoid eviction** by regular VMs!



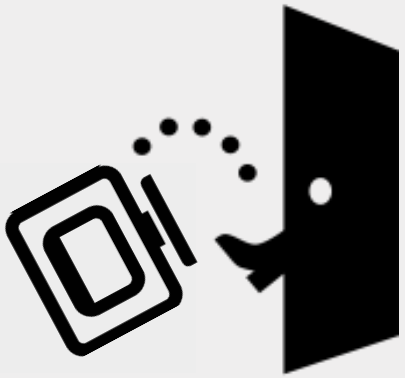
Harvest VMs: Promising for Improved Resource Utilization

Lower Cost

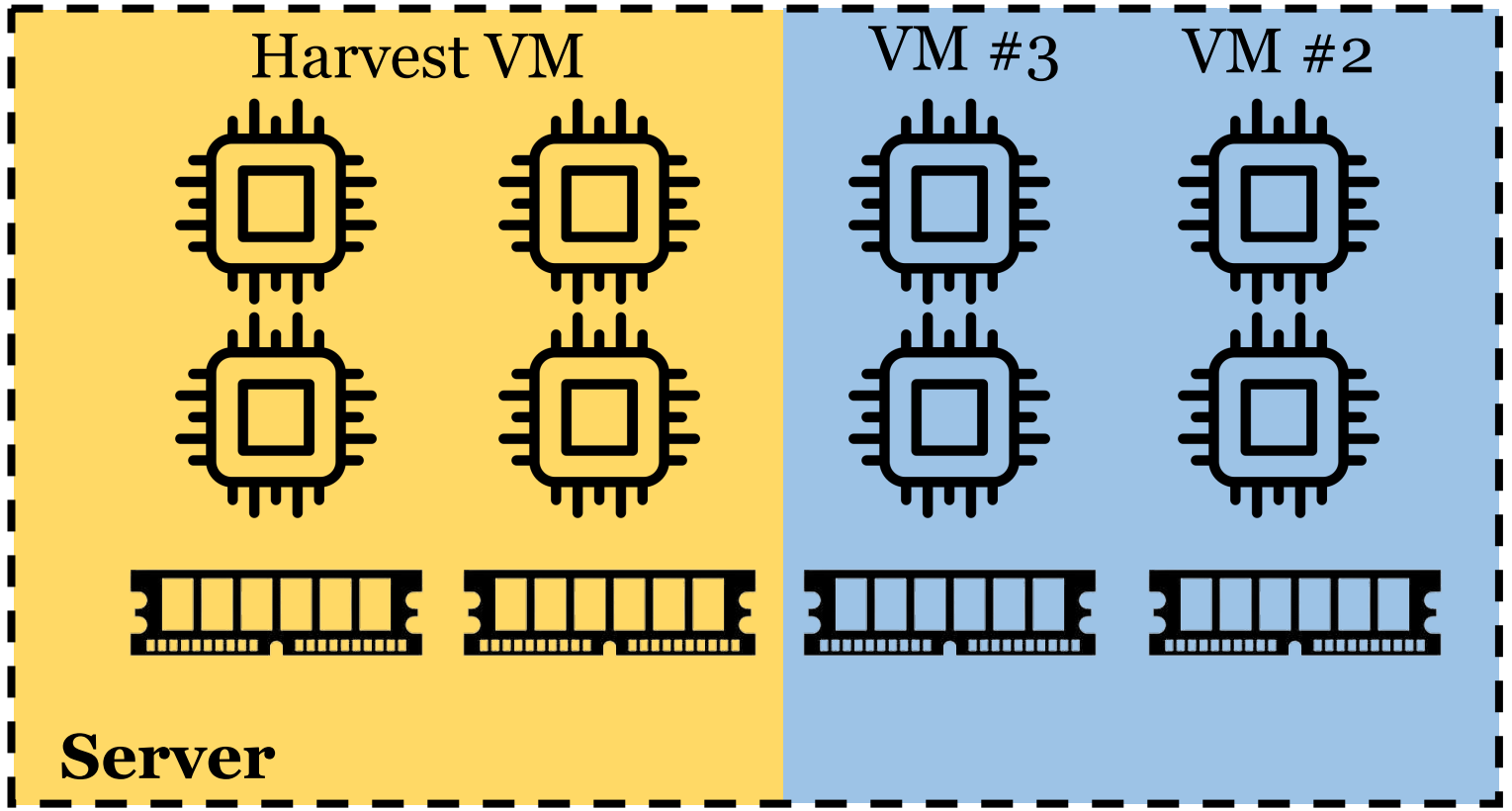


Harvest VMs: Promising for Improved Resource Utilization

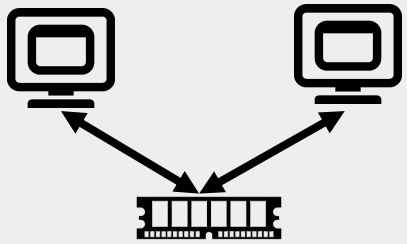
Lower Cost



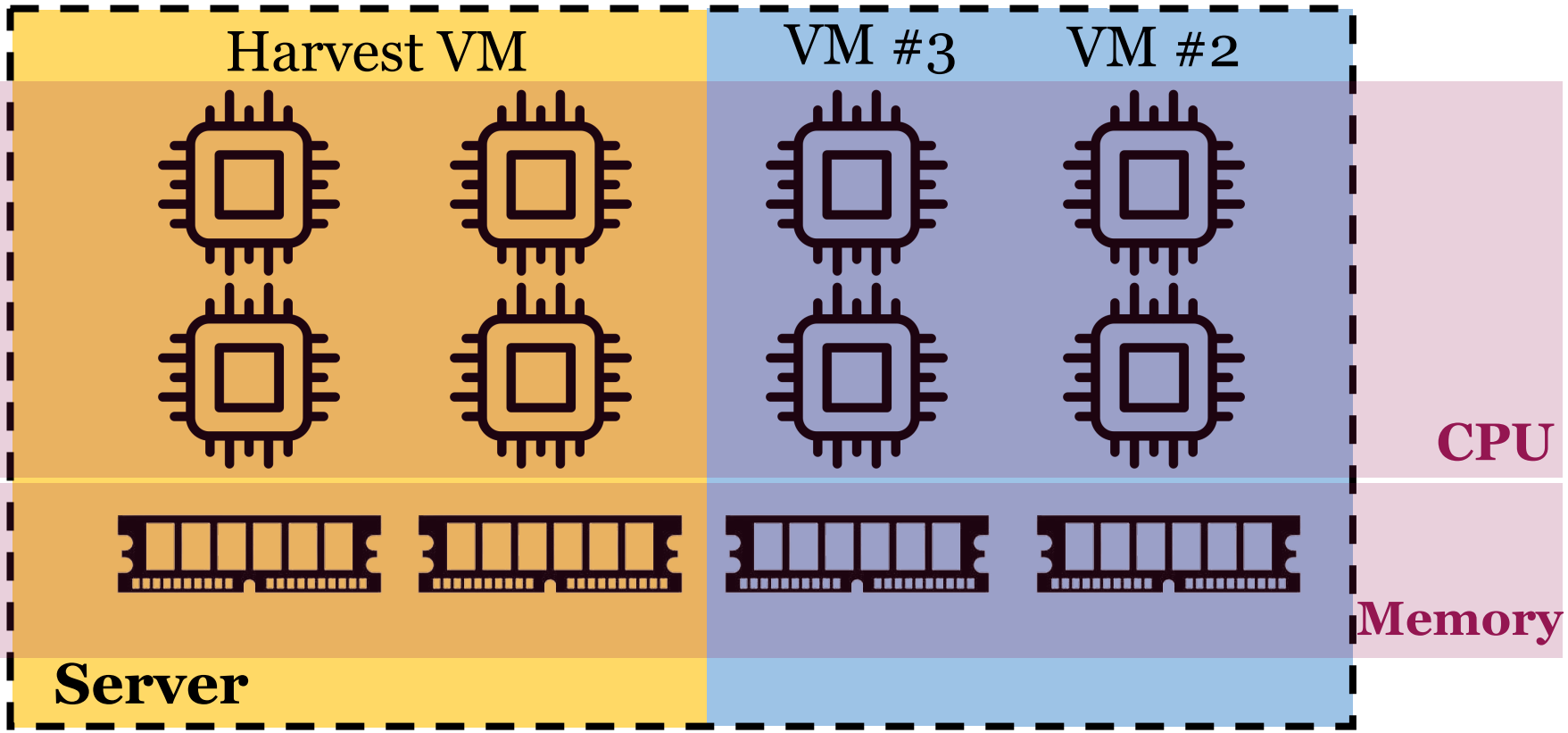
May be Evicted for a Regular VM



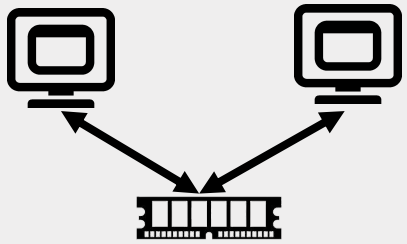
Harvest VMs: Promising for Improved Resource Utilization



Mature Virtualization Support for CPU/Mem

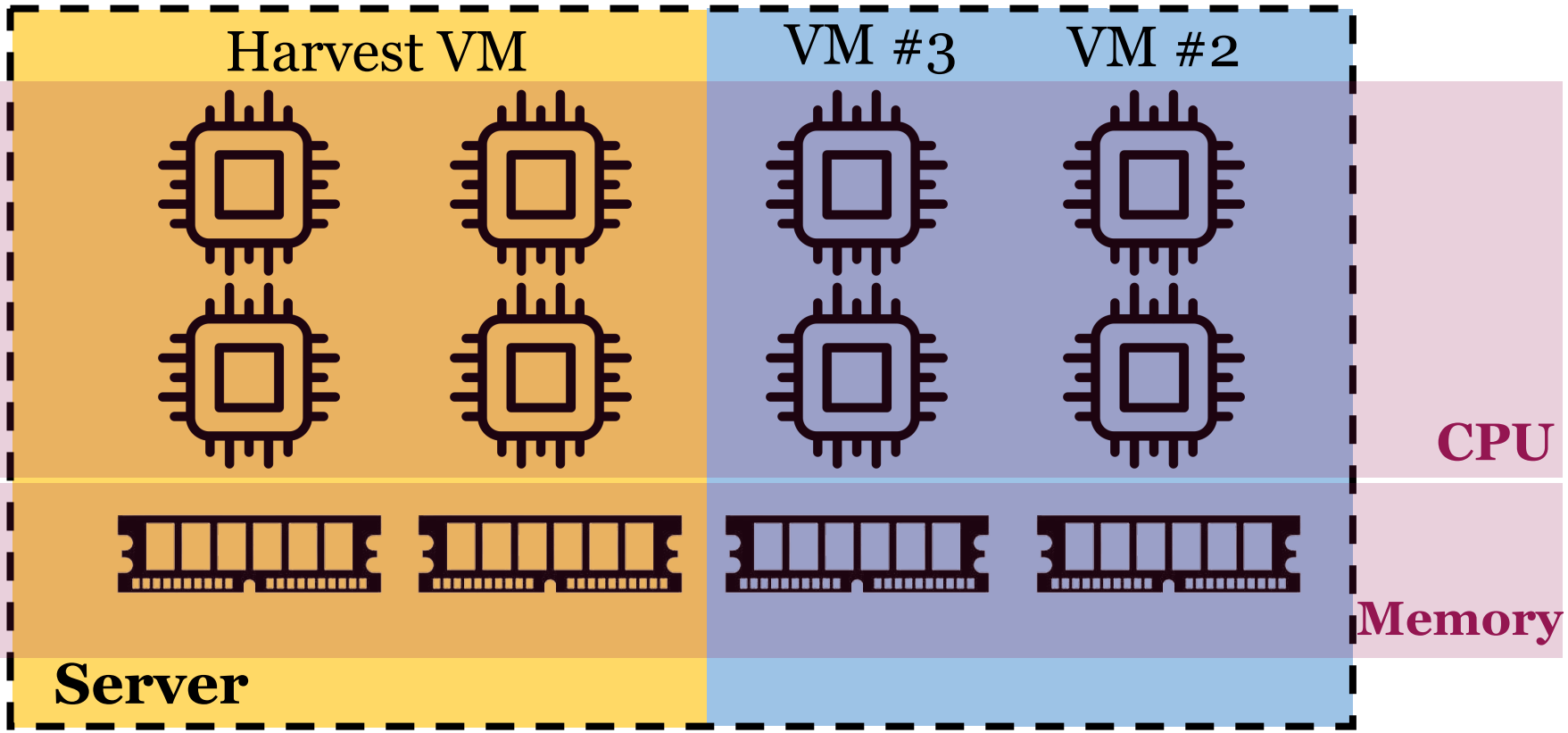


Harvest VMs: Promising for Improved Resource Utilization



Mature Virtualization Support for CPU/Mem

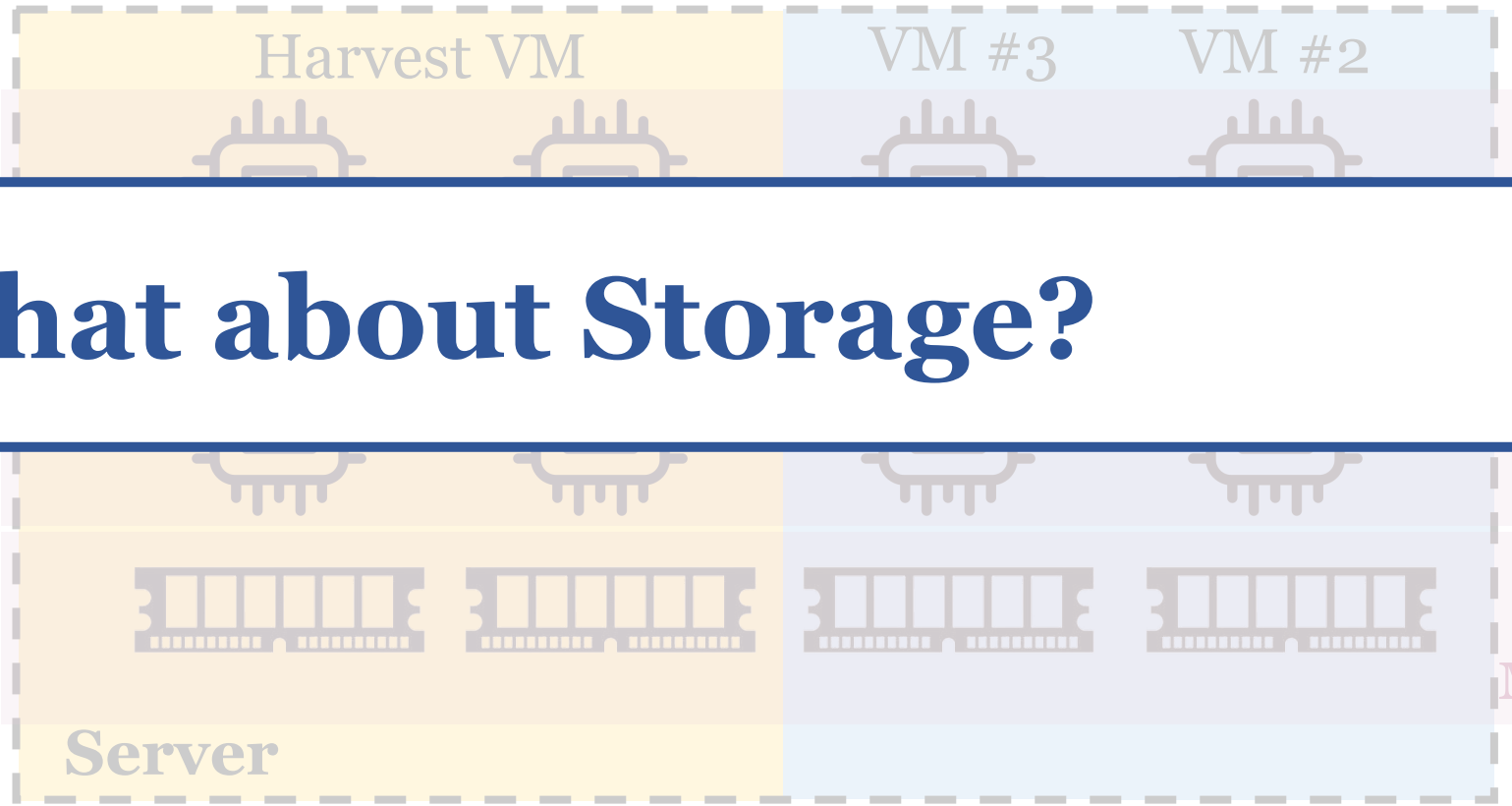
Low Overhead for Resource Scheduling



Harvest VMs: Promising for Improved Resource Utilization

Mature Virtualization

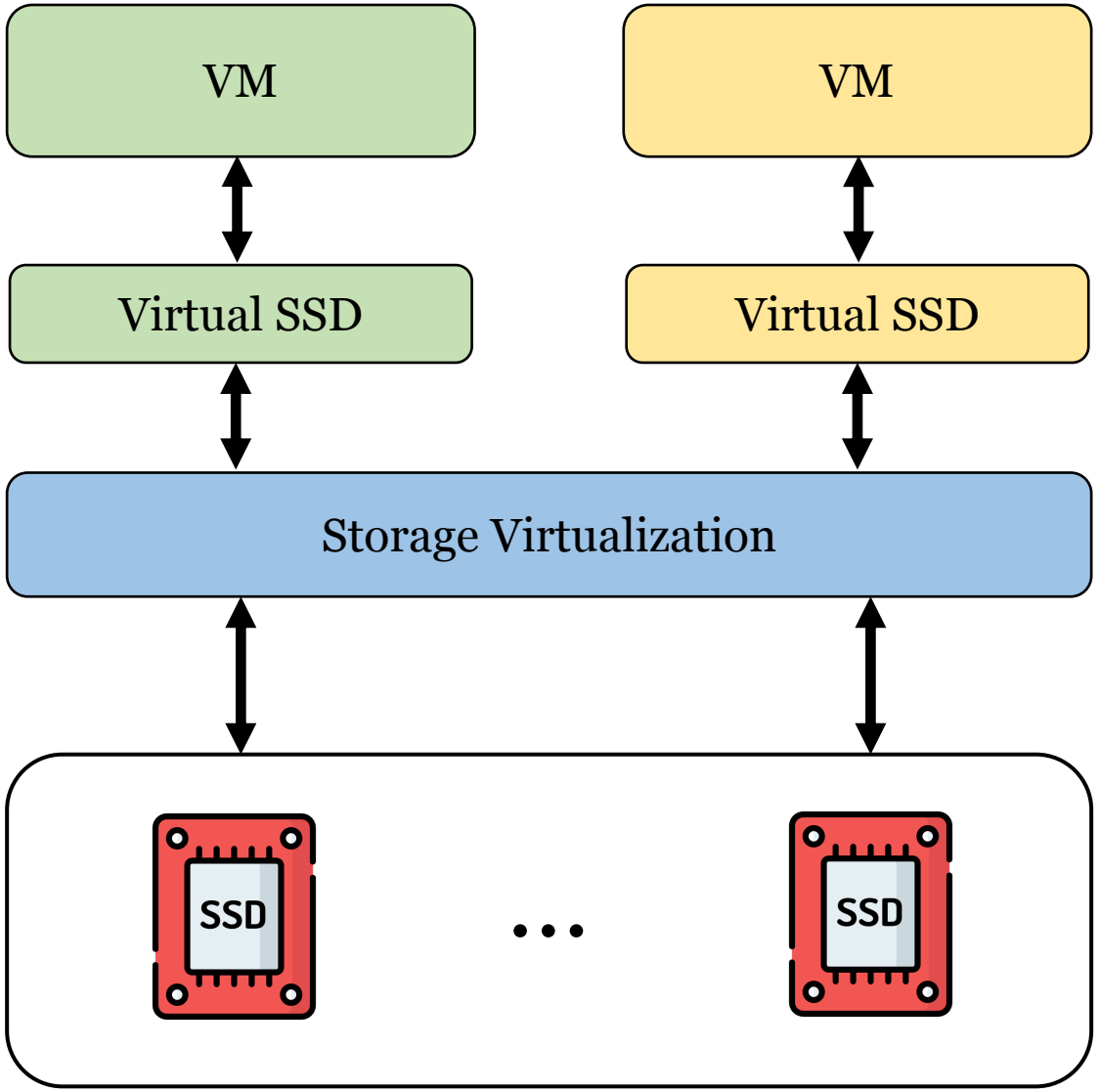
Low Overhead for Resource Scheduling



What about Storage?

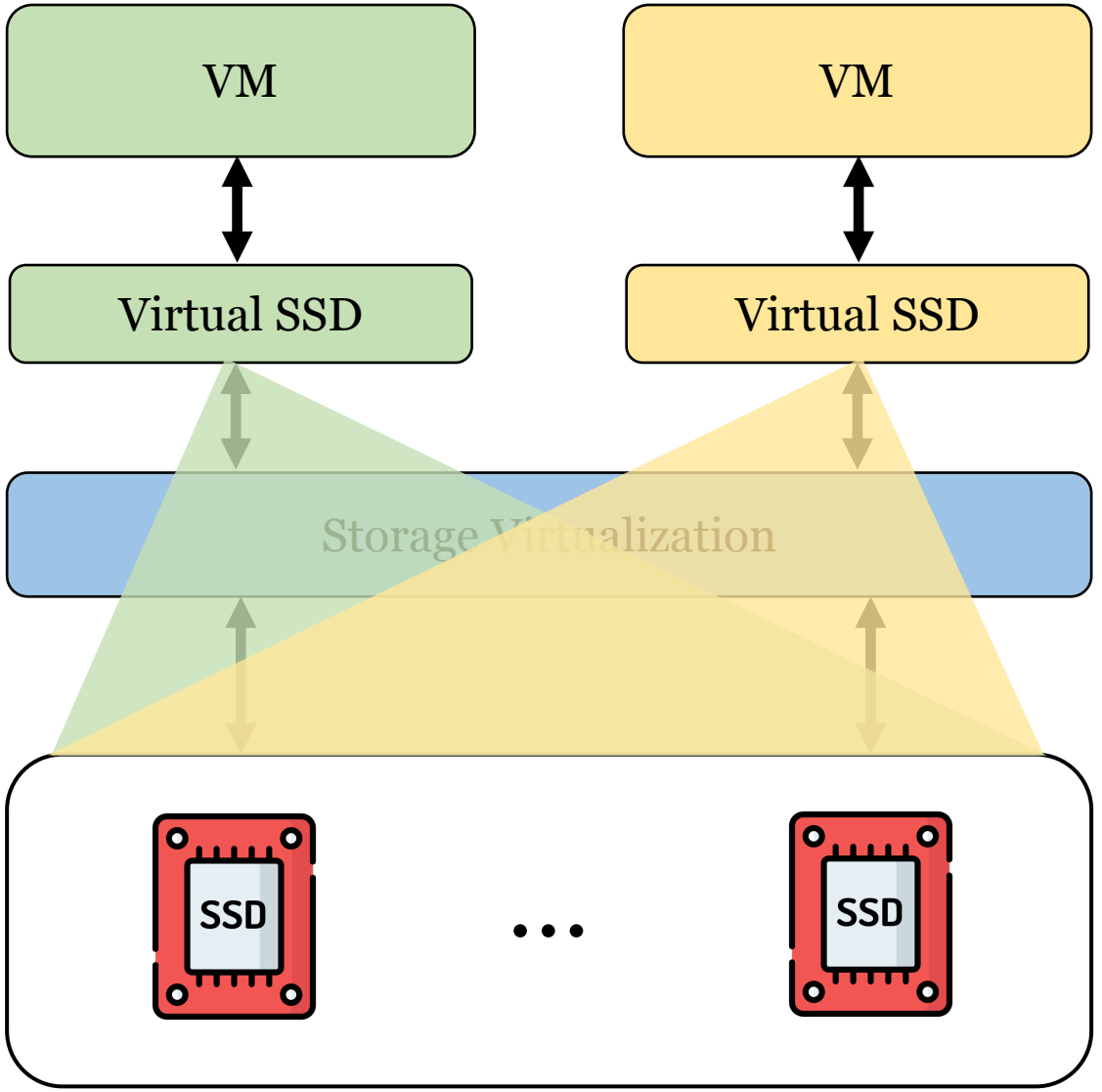
Storage Harvesting is Uniquely Challenging

Insufficient Virtualization Support



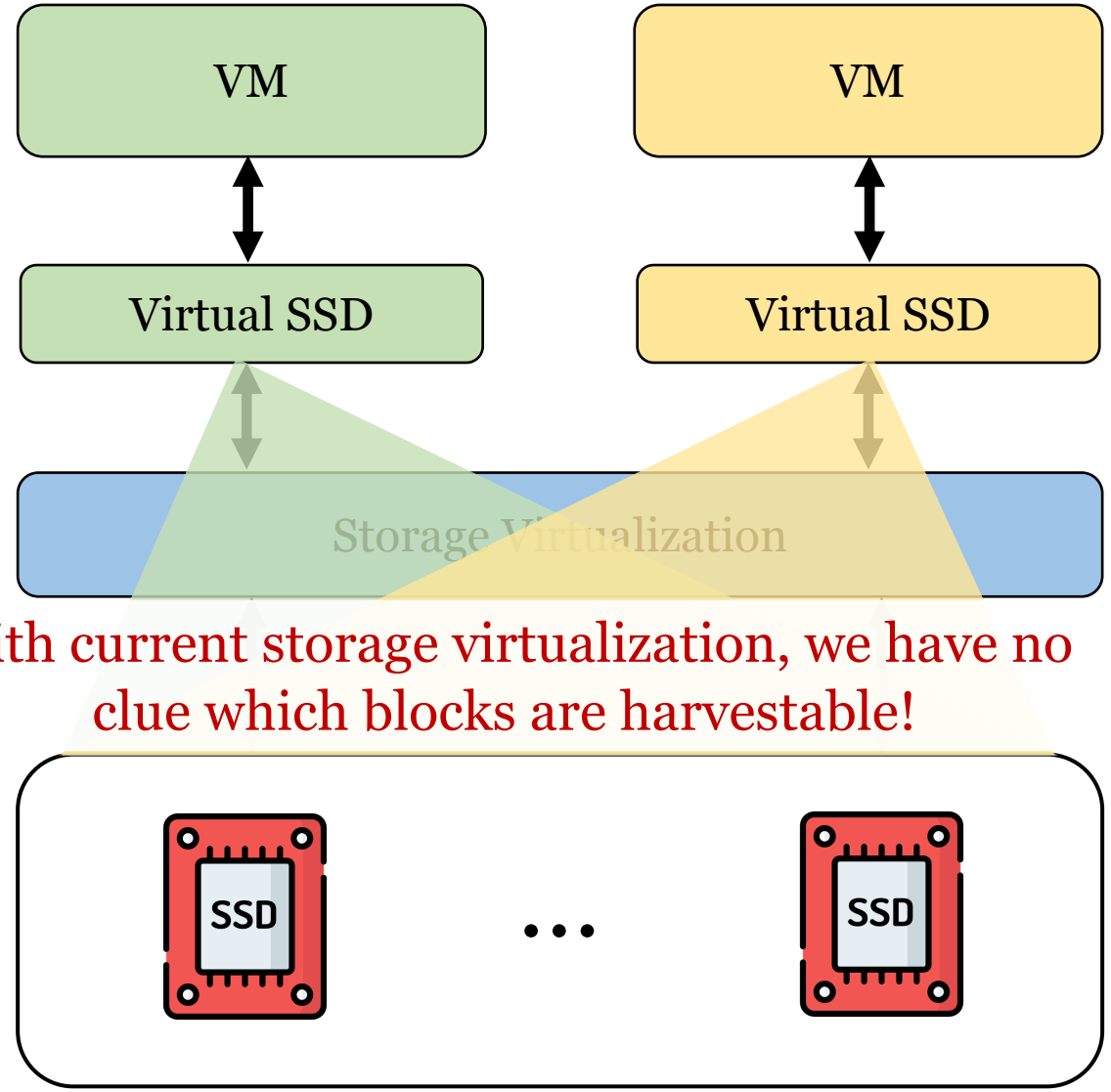
Storage Harvesting is Uniquely Challenging

Insufficient Virtualization Support



Storage Harvesting is Uniquely Challenging

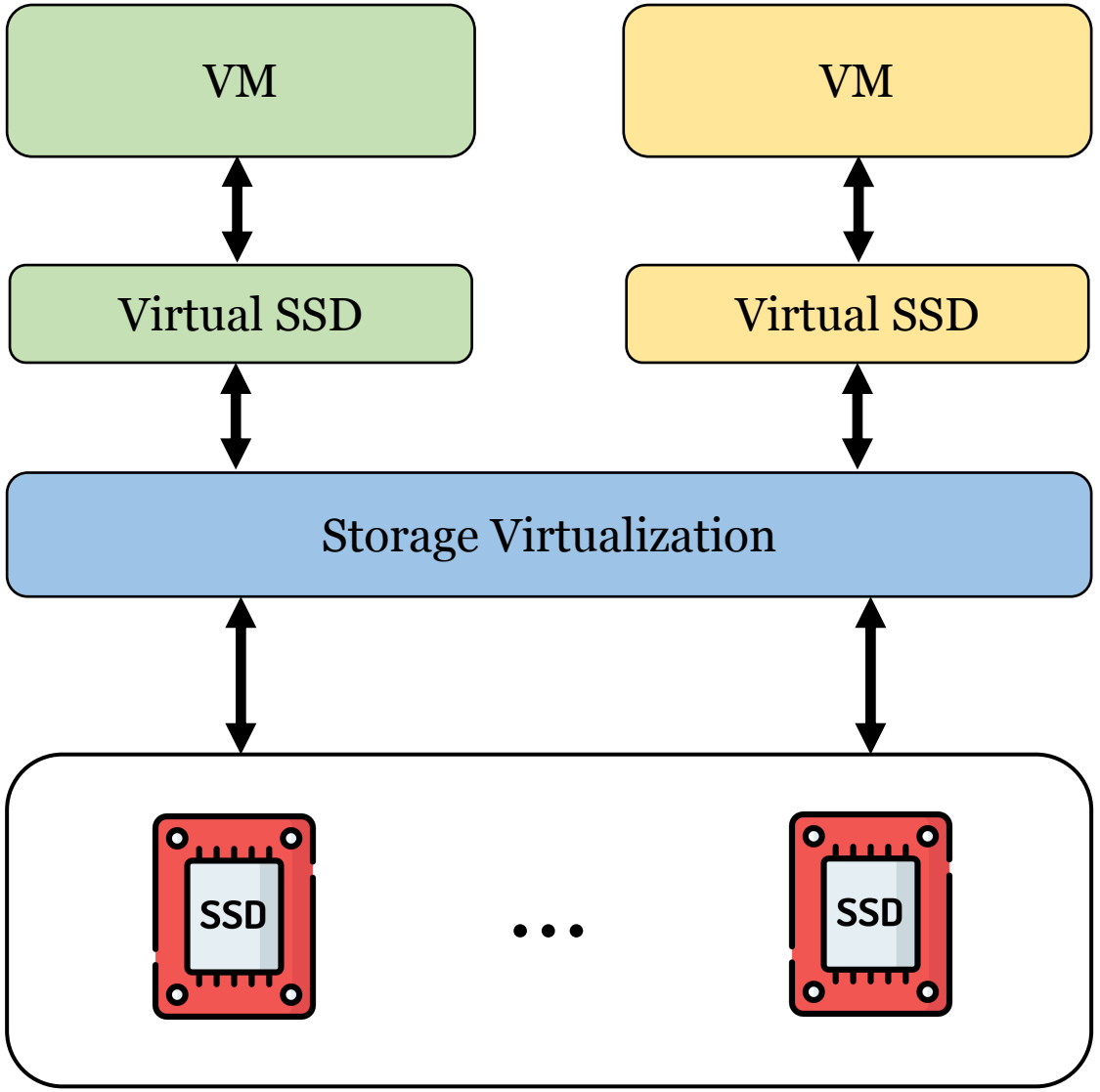
Insufficient Virtualization Support



Storage Harvesting is Uniquely Challenging

Insufficient Virtualization Support

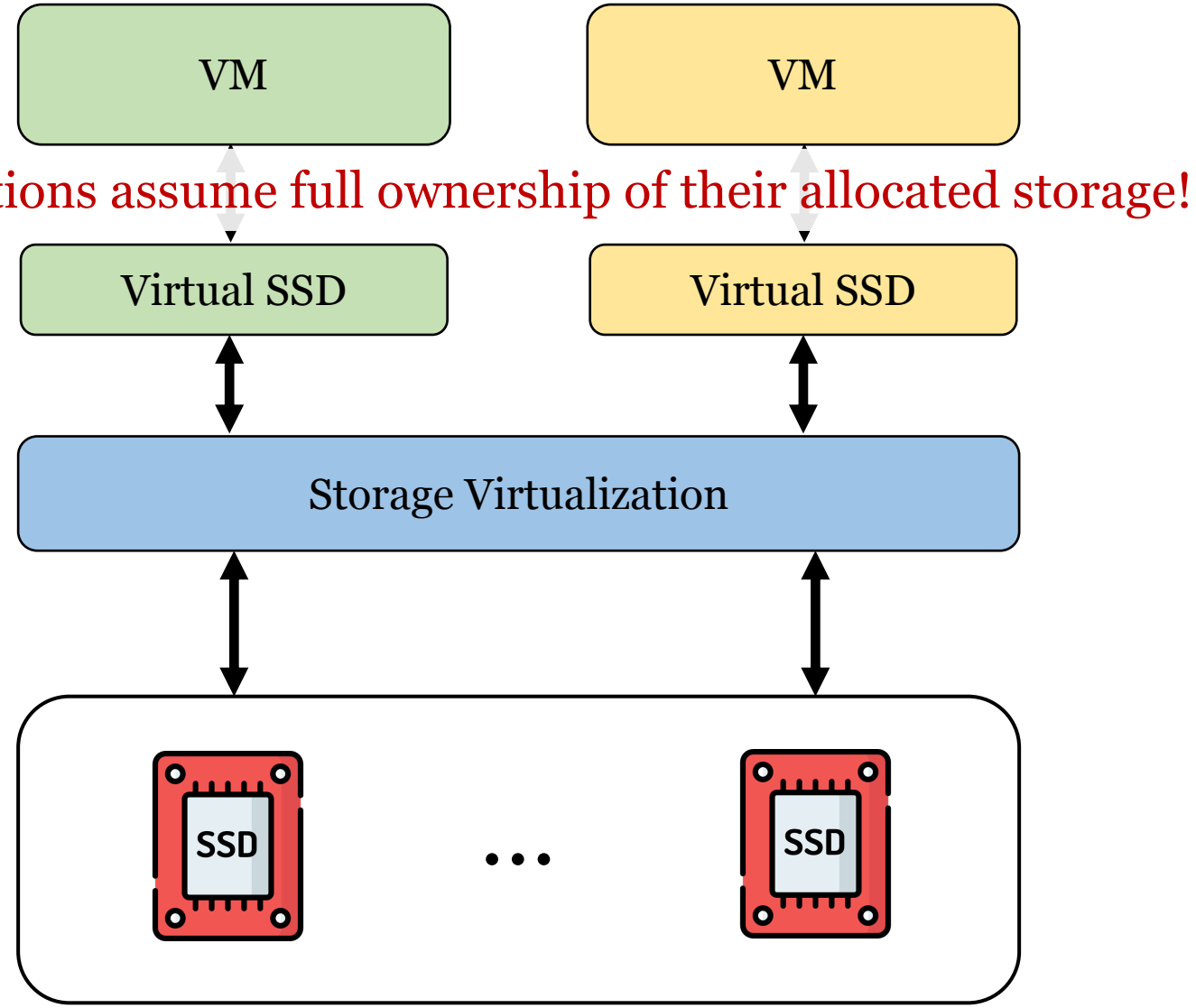
Data Security Concern



Storage Harvesting is Uniquely Challenging

Insufficient Virtualization Support

Data Security Concern

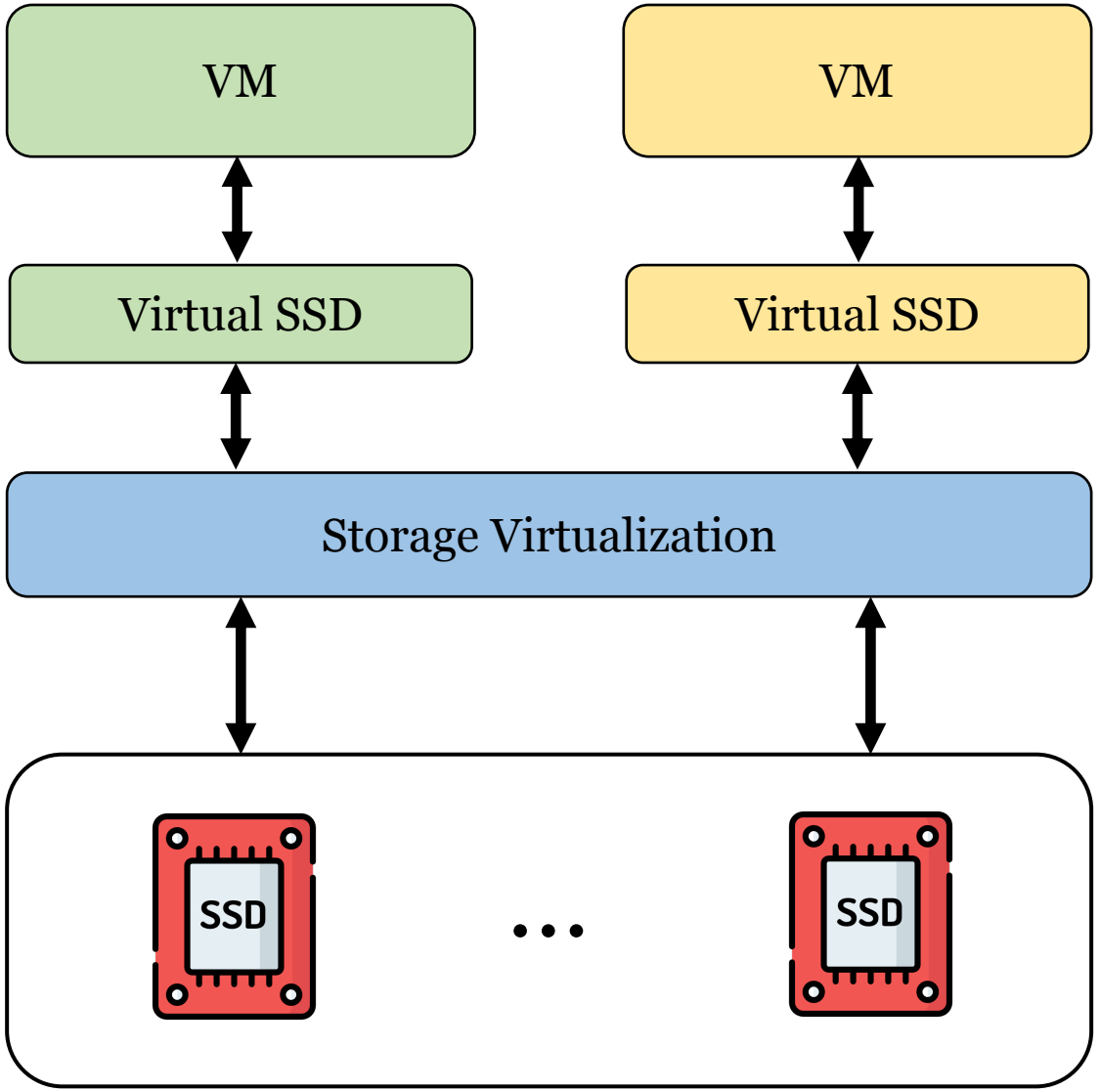


Storage Harvesting is Uniquely Challenging

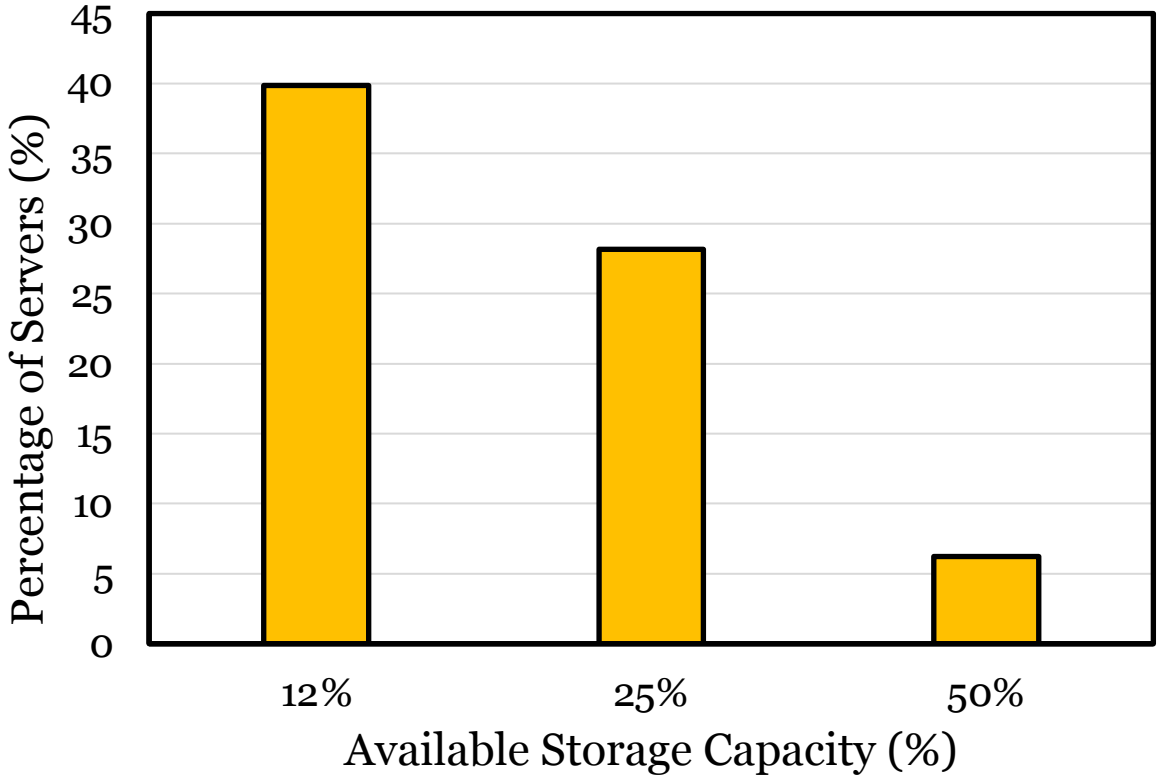
Insufficient Virtualization Support

Data Security Concern

Storage Performance Interference

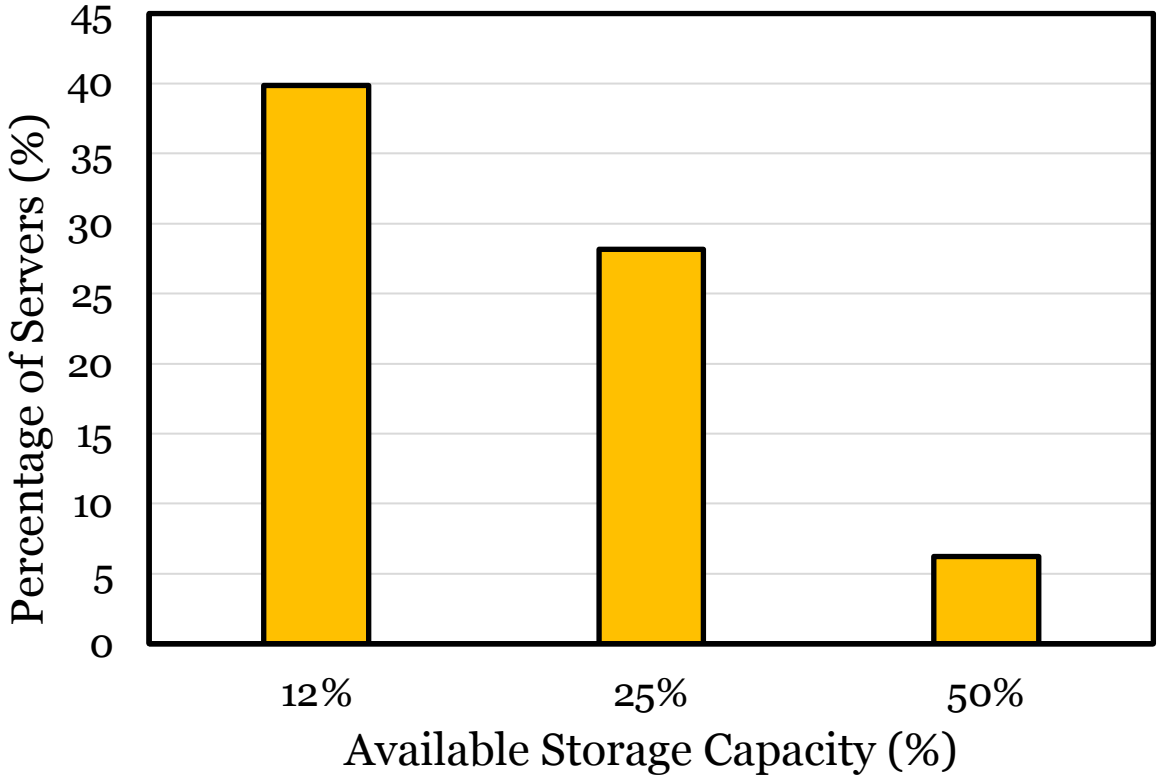


Is Storage Harvesting Worth It?



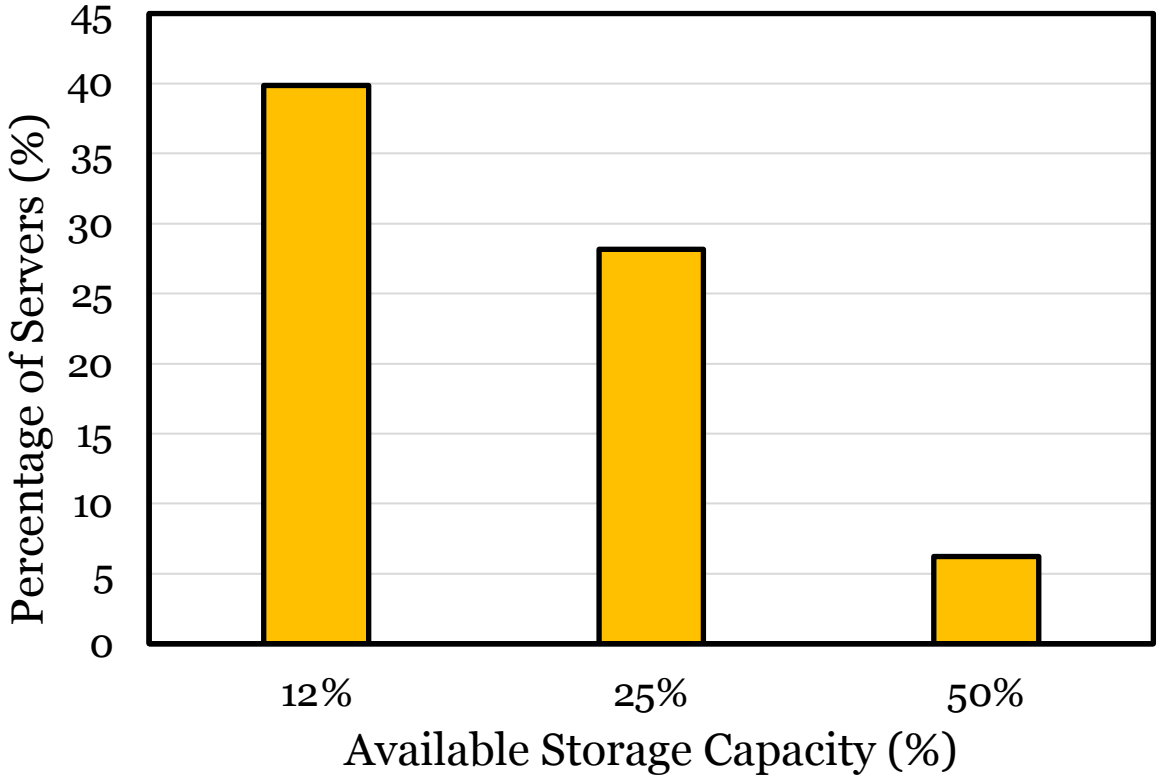
Is Storage Harvesting Worth It?

40% of servers have harvestable storage capacity



Is Storage Harvesting Worth It?

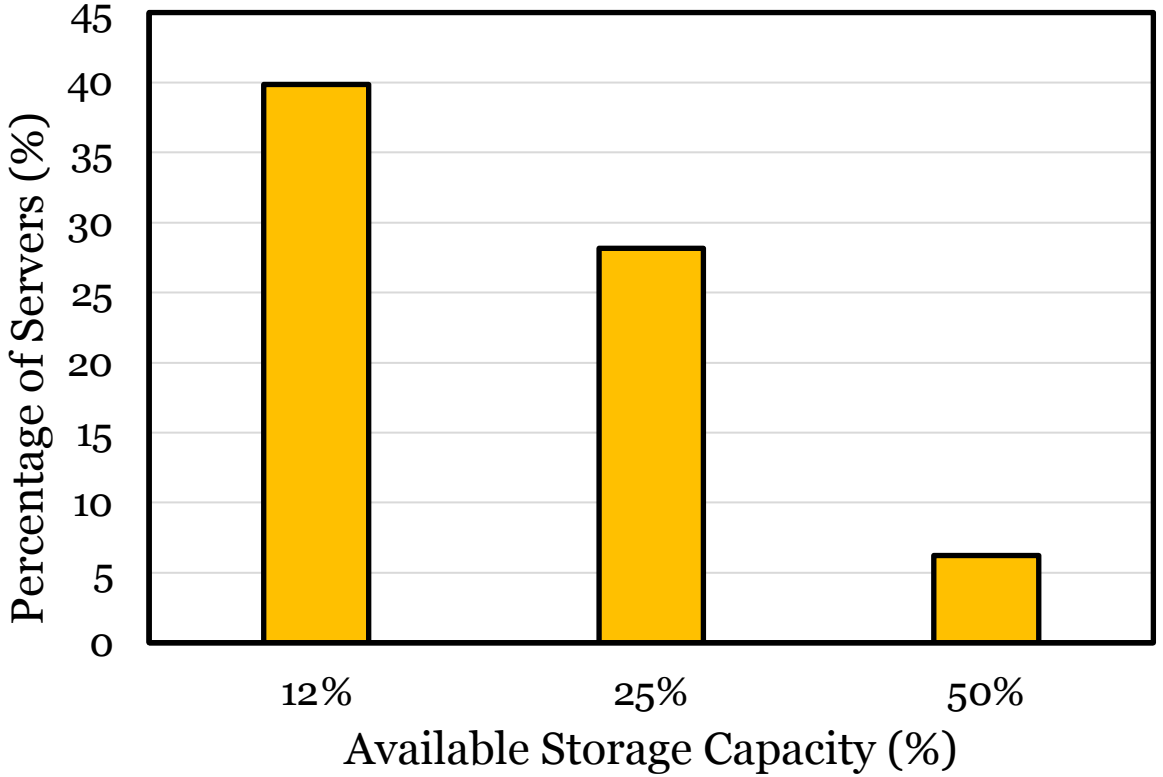
40% of servers have harvestable storage capacity

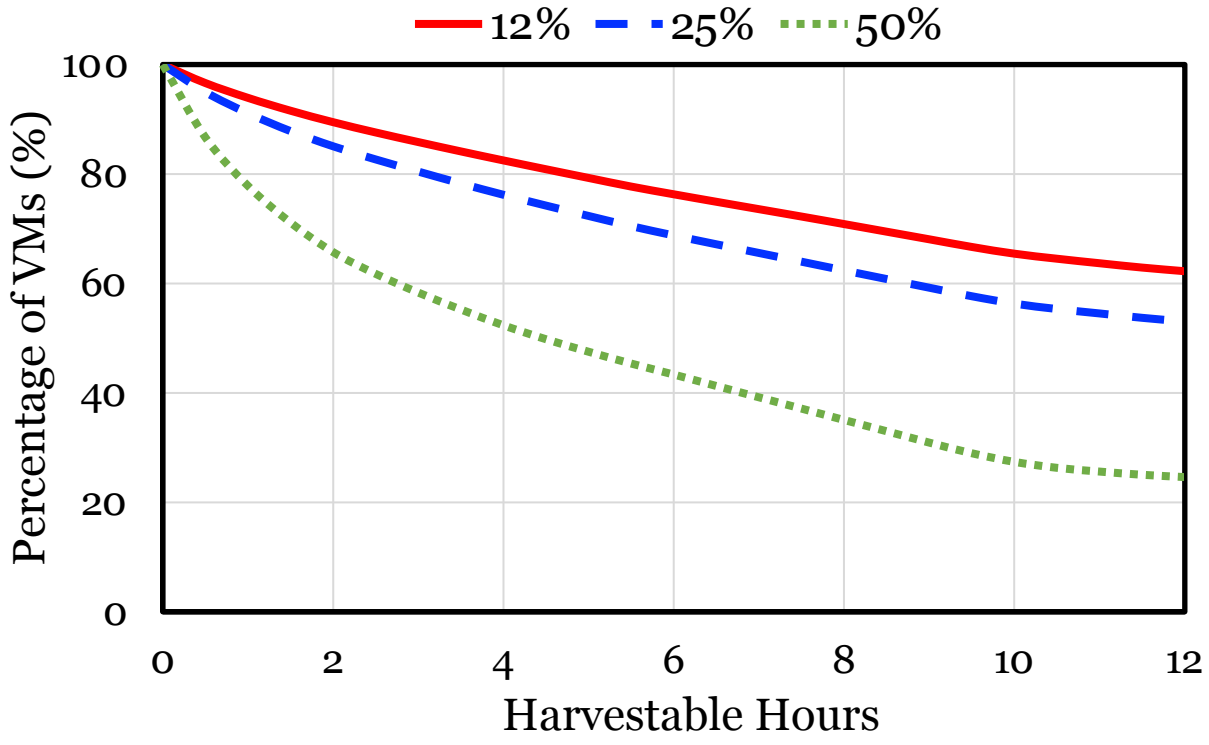


Over 6% of servers have **50% of their storage available**

Is Storage Harvesting Worth It?

40% of servers have harvestable storage capacity

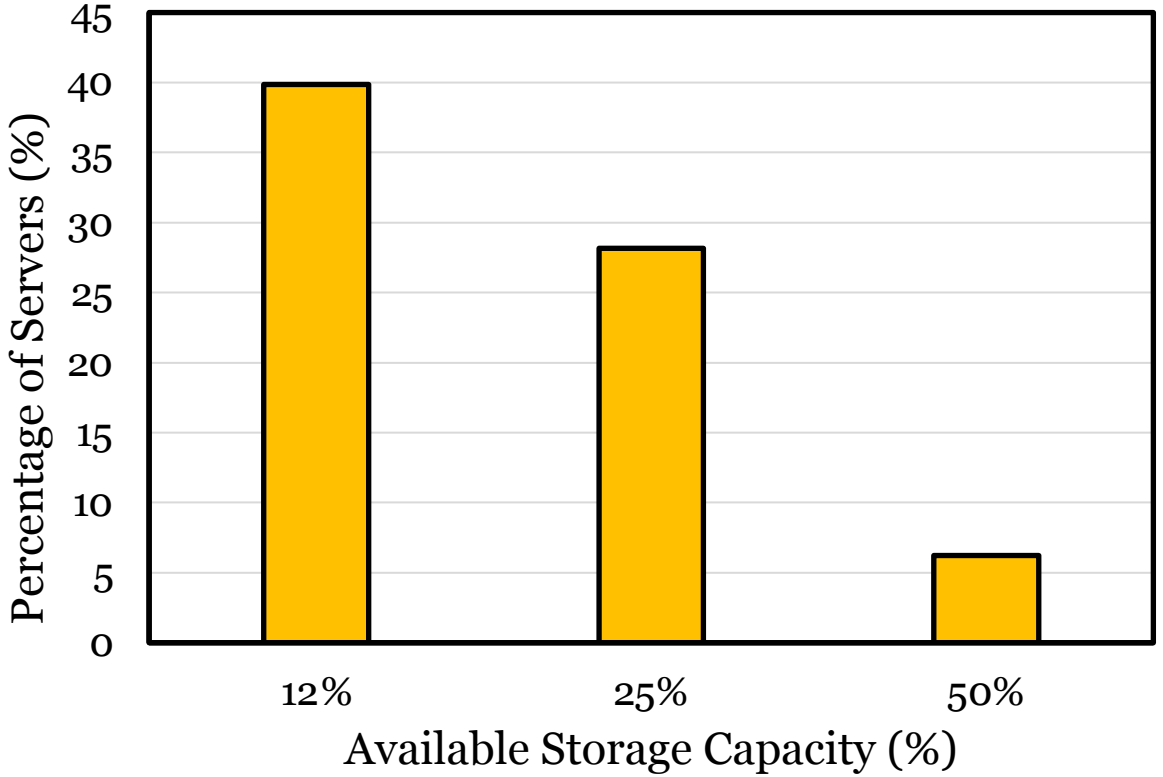




Over 6% of servers have **50% of their storage available**

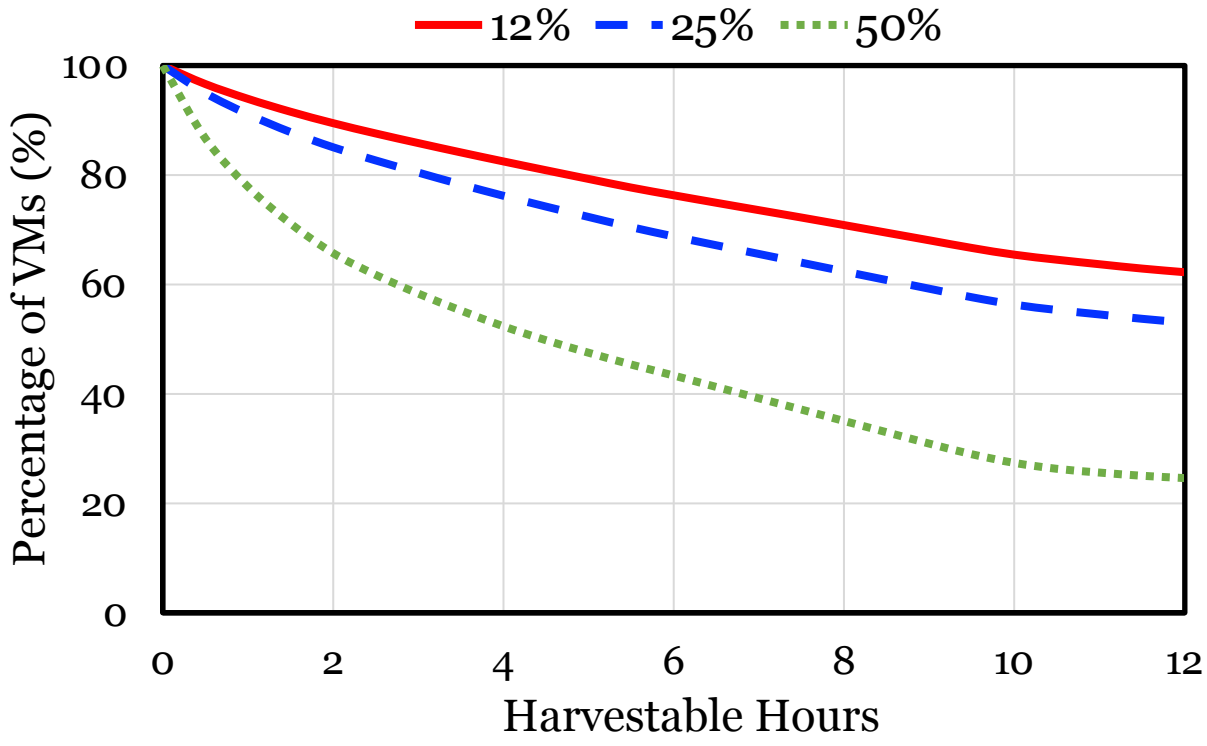
Is Storage Harvesting Worth It?

40% of servers have harvestable storage capacity



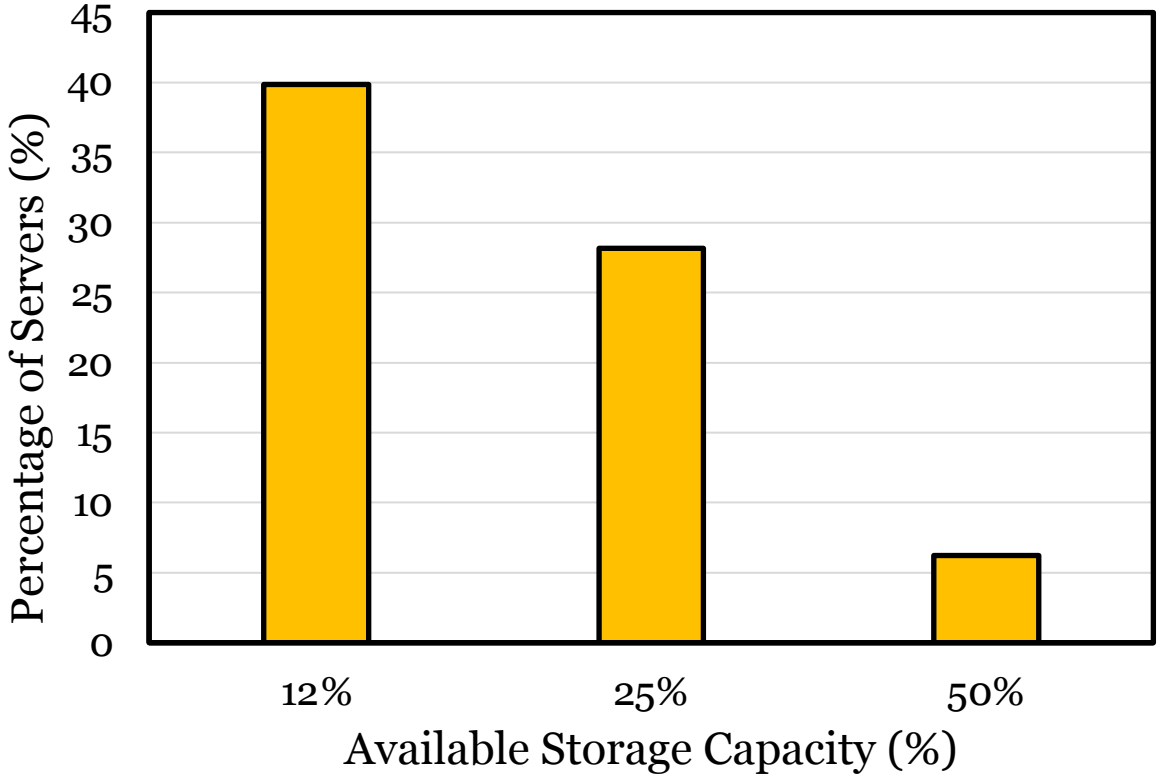
Over 6% of servers have **50% of their storage available**

Over 60% of harvestable storage **lasts over 12 hours**



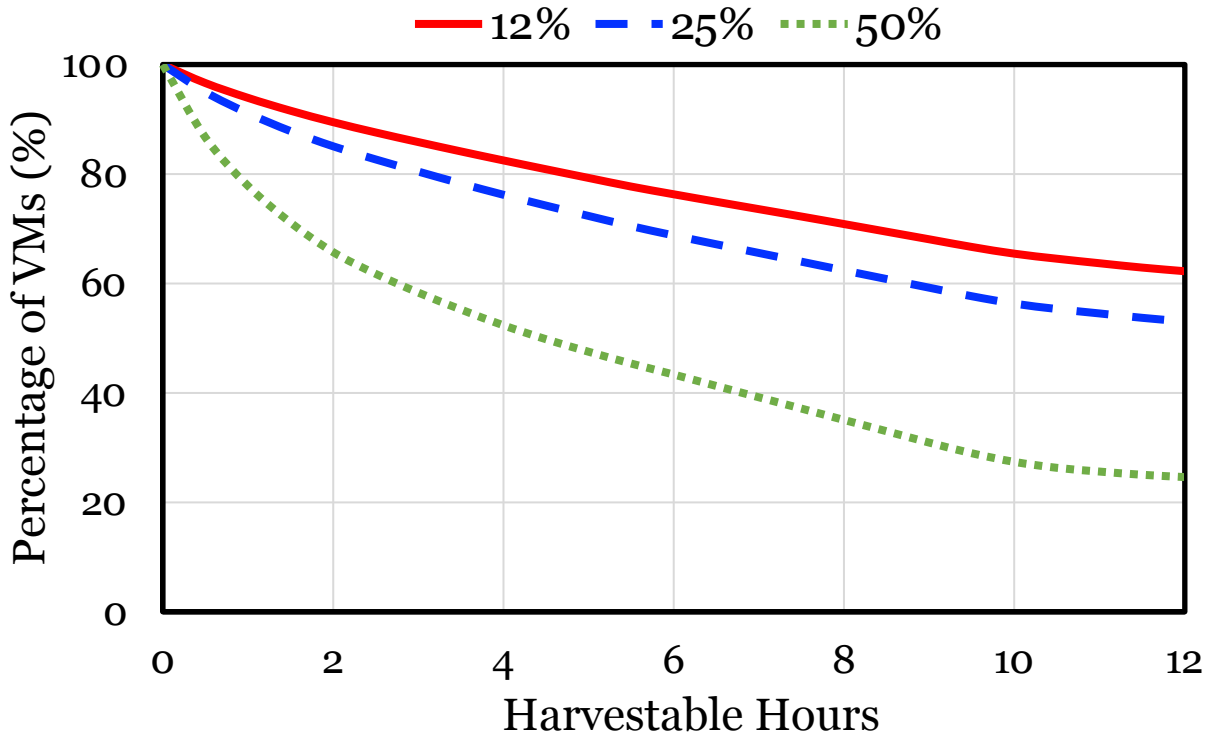
Is Storage Harvesting Worth It?

40% of servers have harvestable storage capacity



Over 6% of servers have **50% of their storage available**

Over 60% of harvestable storage **lasts over 12 hours**

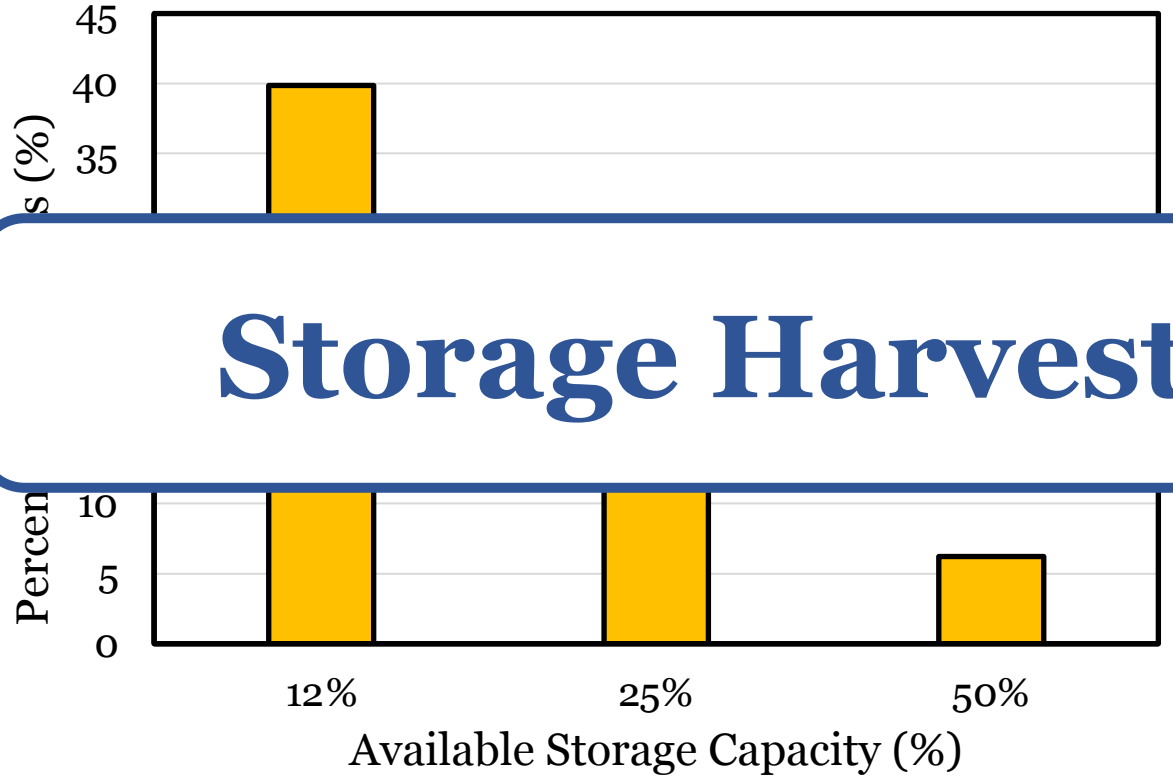


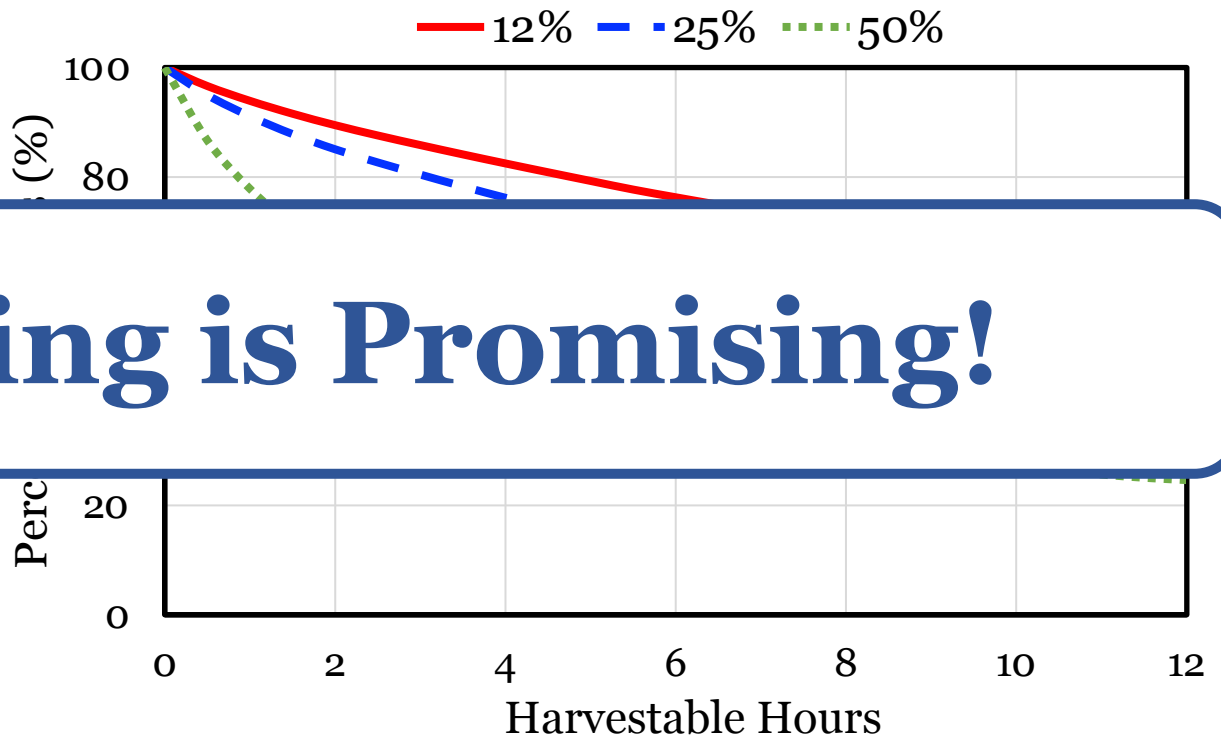
Harvesting of all sizes shows **low chance of eviction**

Is Storage Harvesting Worth It?

40% of servers have harvestable storage capacity

Over 60% of harvestable storage **lasts over 12 hours**





Storage Harvesting is Promising!

Over 6% of servers have **50% of their storage available**

Harvesting of all sizes shows **low chance of eviction**

BlockFlex: Enabling Storage Harvesting in Modern Cloud

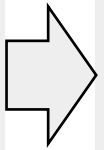
Insufficient Virtualization Support

Data Security Concern

Storage Performance Interference

BlockFlex: Enabling Storage Harvesting in Modern Cloud

Insufficient Virtualization Support



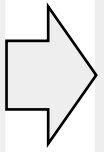
Extend Storage Virtualization with **Ghost vSSD Abstraction**

Data Security Concern

Storage Performance Interference

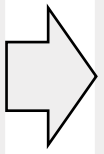
BlockFlex: Enabling Storage Harvesting in Modern Cloud

Insufficient Virtualization Support



Extend Storage Virtualization with **Ghost vSSD Abstraction**

Data Security Concern



Enable Fine-Grained **Hardware Isolation** and **Block Erasing**

Storage Performance Interference

BlockFlex: Enabling Storage Harvesting in Modern Cloud

Insufficient Virtualization Support

Extend Storage Virtualization with **Ghost vSSD Abstraction**

Data Security Concern

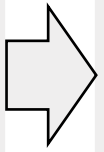
Enable Fine-Grained **Hardware Isolation** and **Block Erasing**

Storage Performance Interference

Minimize the **Impact of Harvesting** on Regular VMs

BlockFlex: Enabling Storage Harvesting in Modern Cloud

Insufficient Virtualization Support



Extend Storage Virtualization with **Ghost vSSD Abstraction**

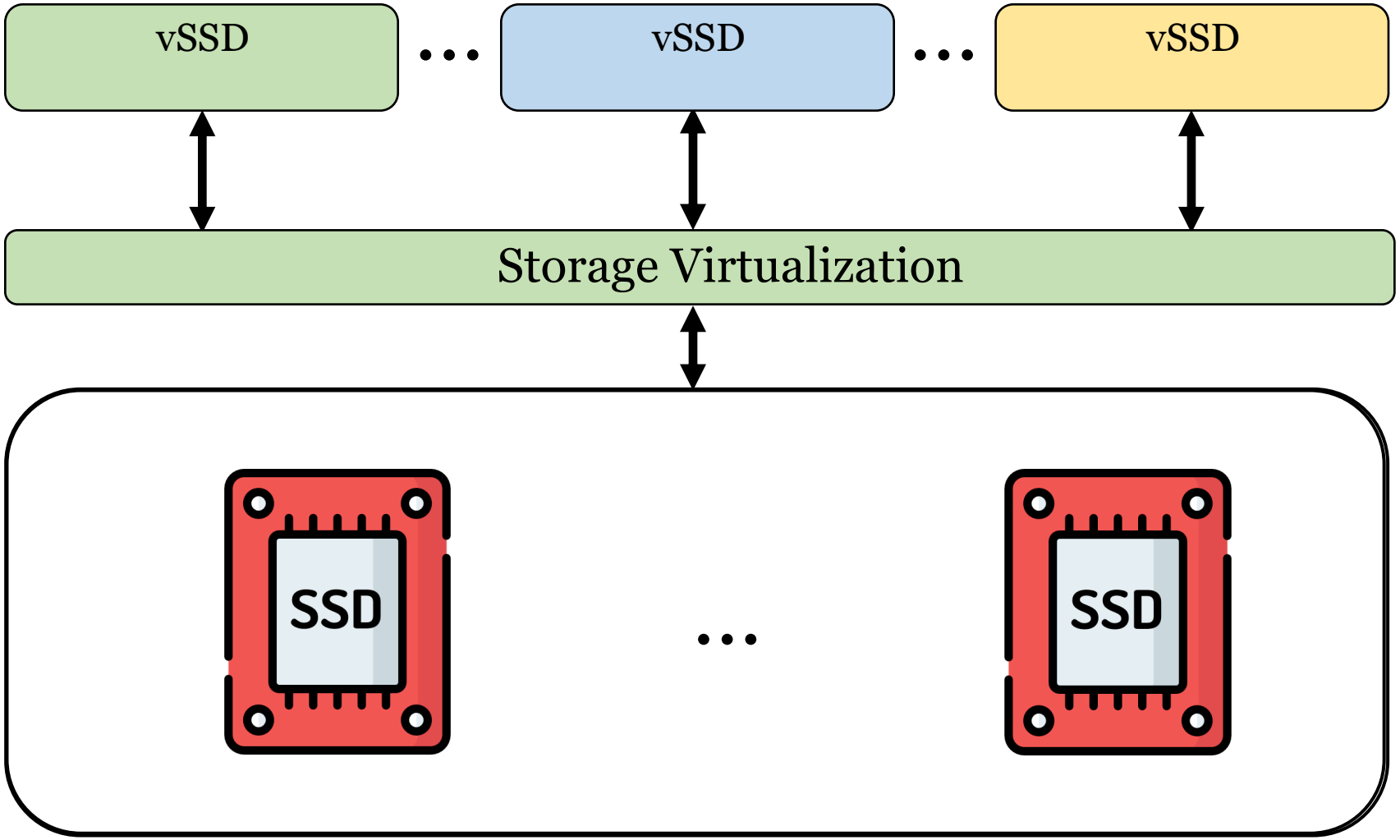
Data Security Concern

Enable Fine-Grained **Hardware Isolation** and **Block Erasing**

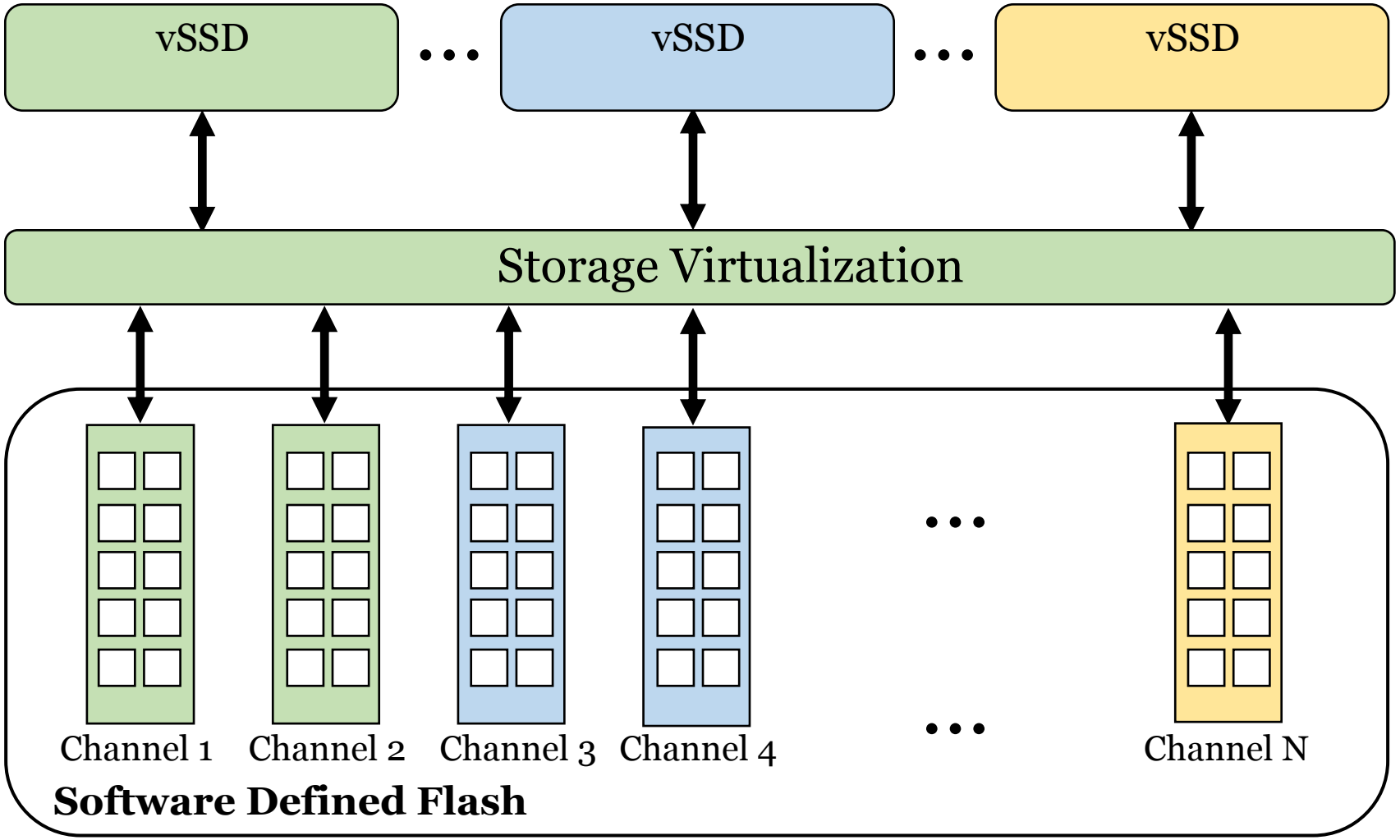
Storage Performance Interference

Minimize the **Impact of Harvesting** on Regular VMs

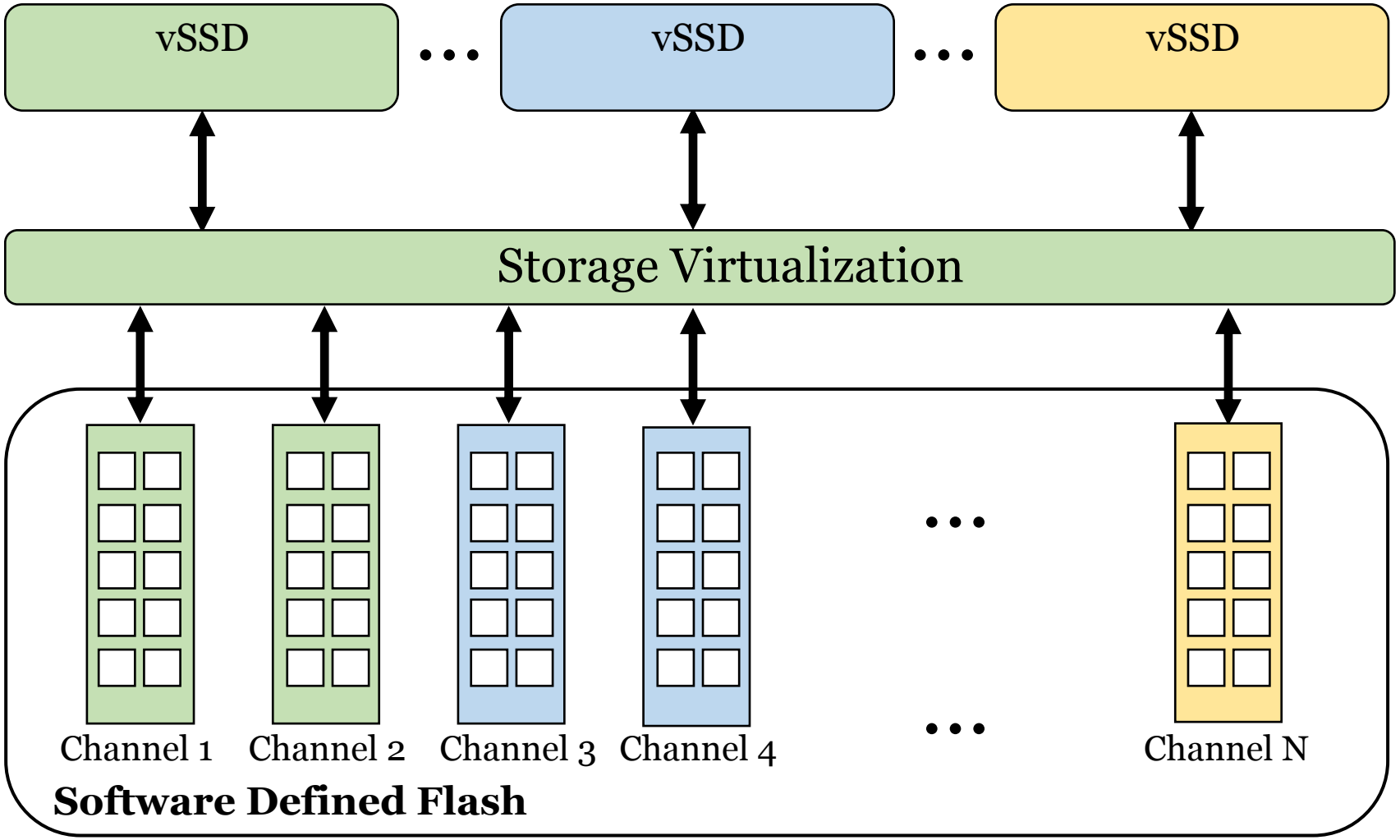
Enabling Storage Harvesting with Software-Defined Flash

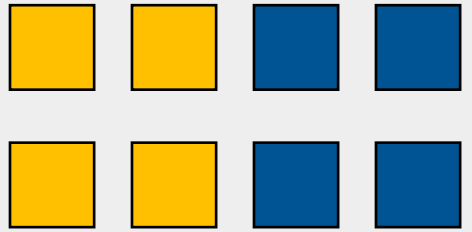


Enabling Storage Harvesting with Software-Defined Flash



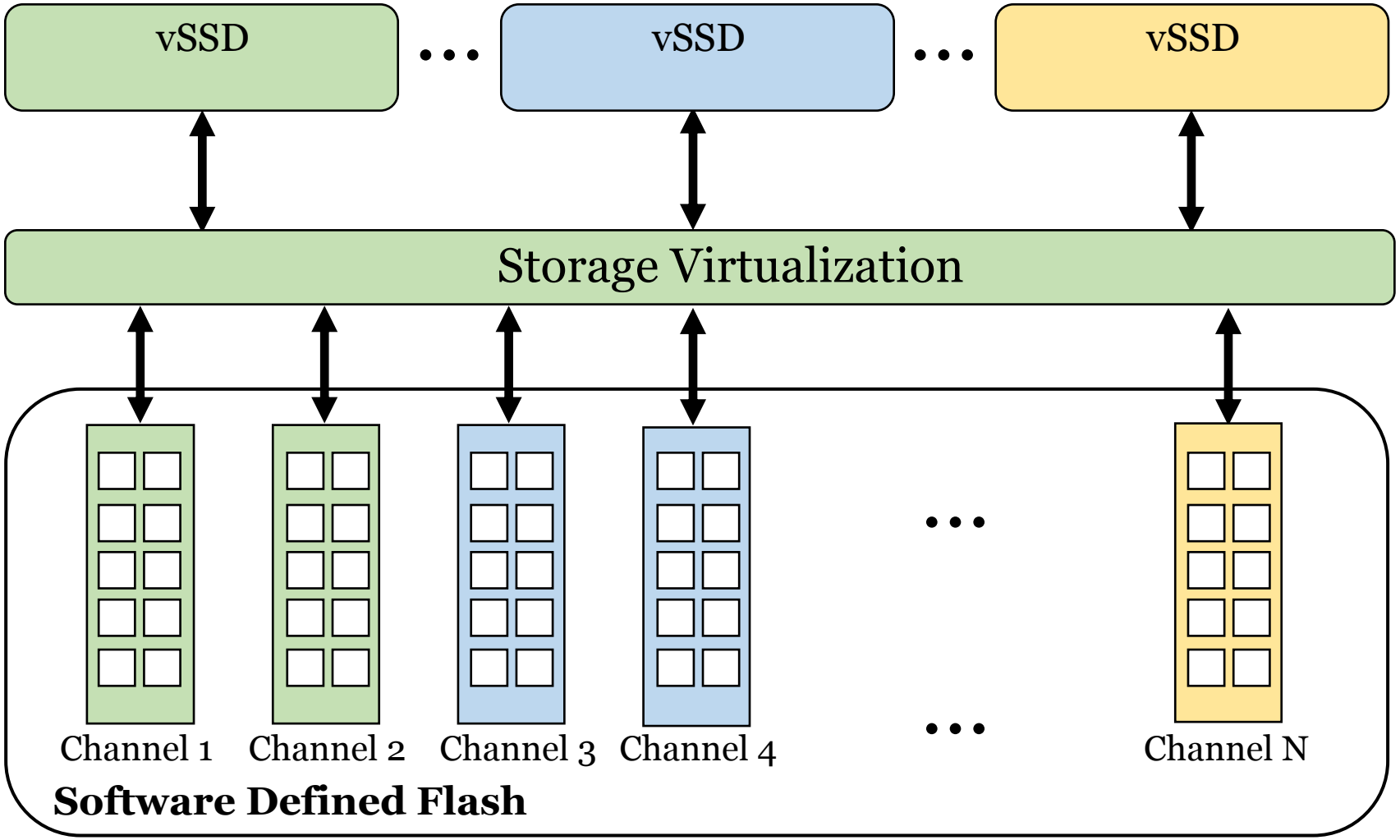
Enabling Storage Harvesting with Software-Defined Flash

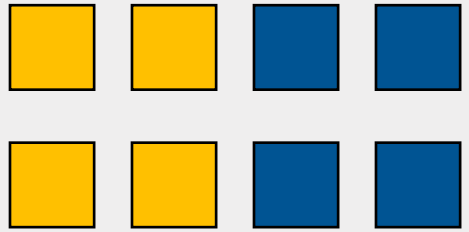




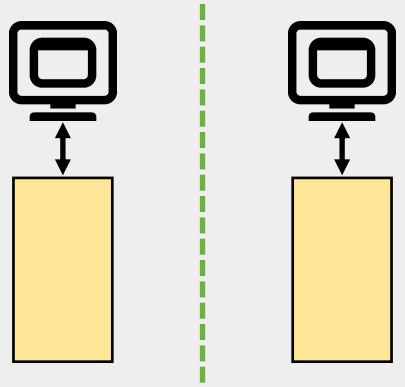
Flexibility for
block management
and **state tracking**

Enabling Storage Harvesting with Software-Defined Flash



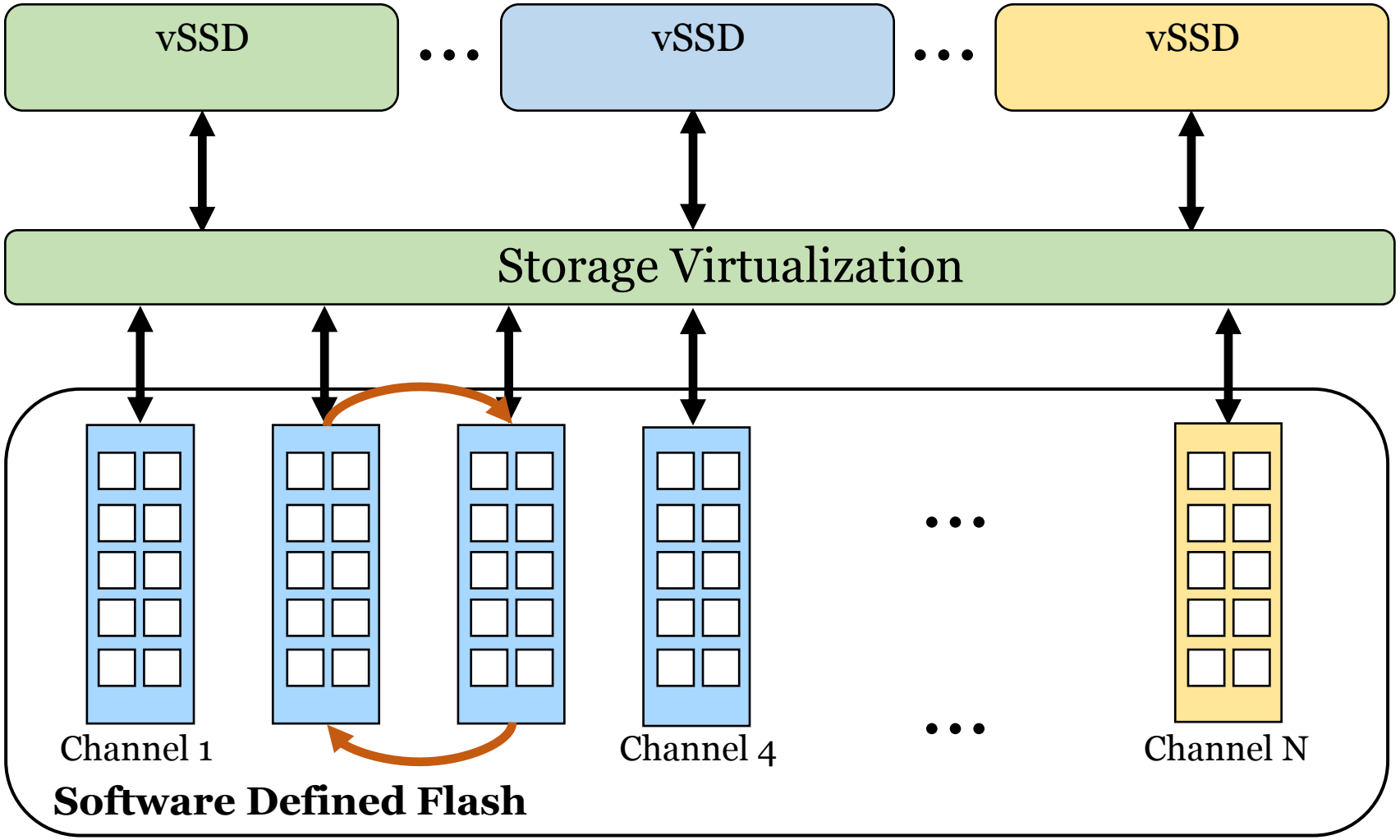


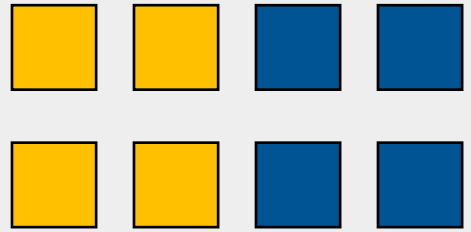
Flexibility for **block management** and **state tracking**



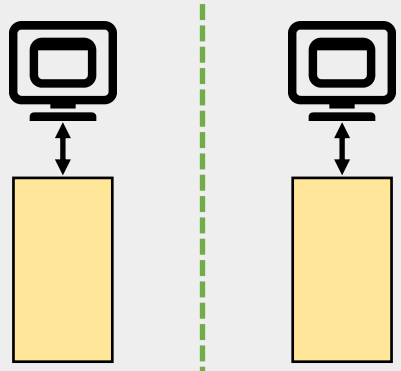
Hardware isolation between vSSDs

Enabling Storage Harvesting with Software-Defined Flash



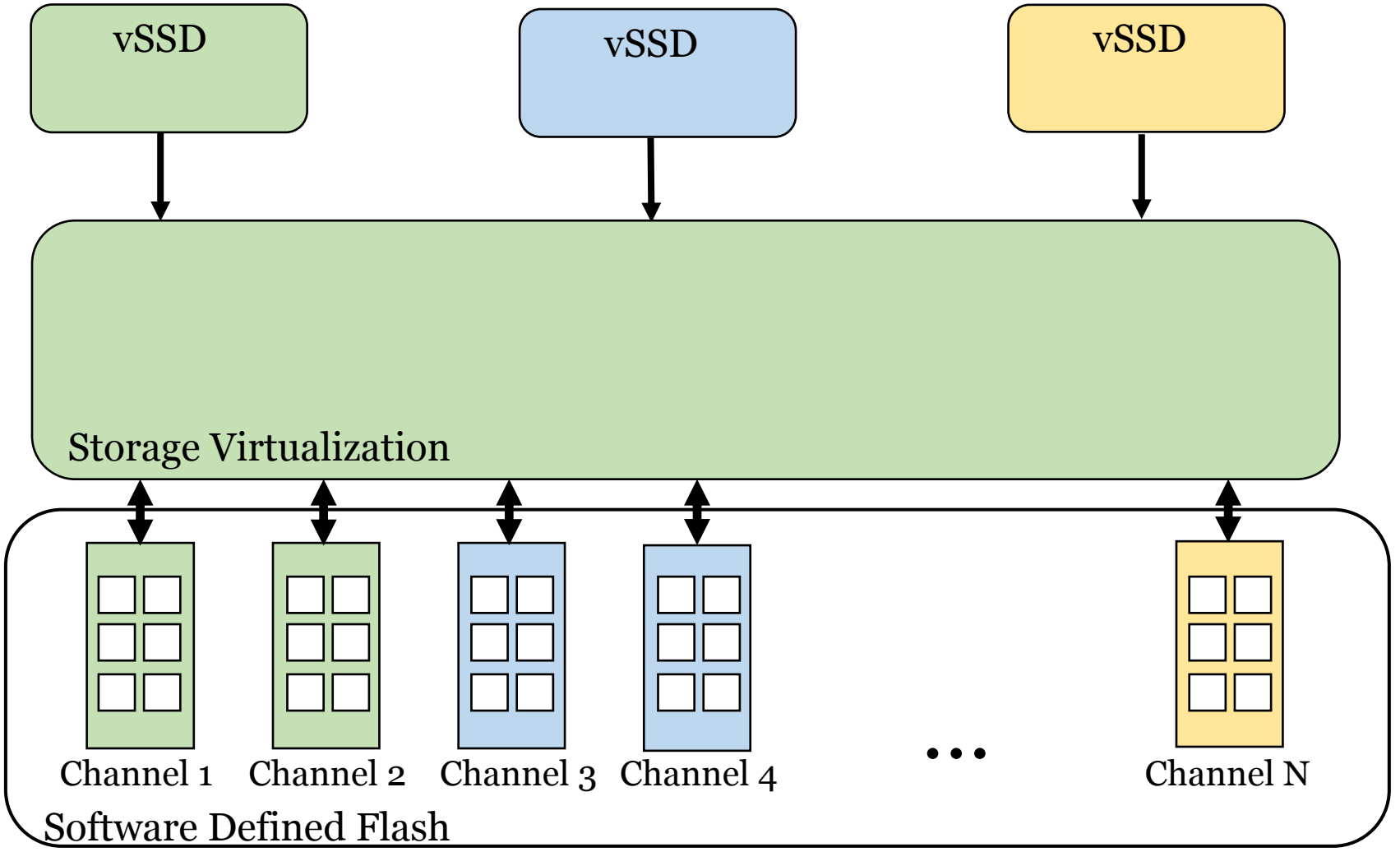


Flexibility for **block management** and **state tracking**

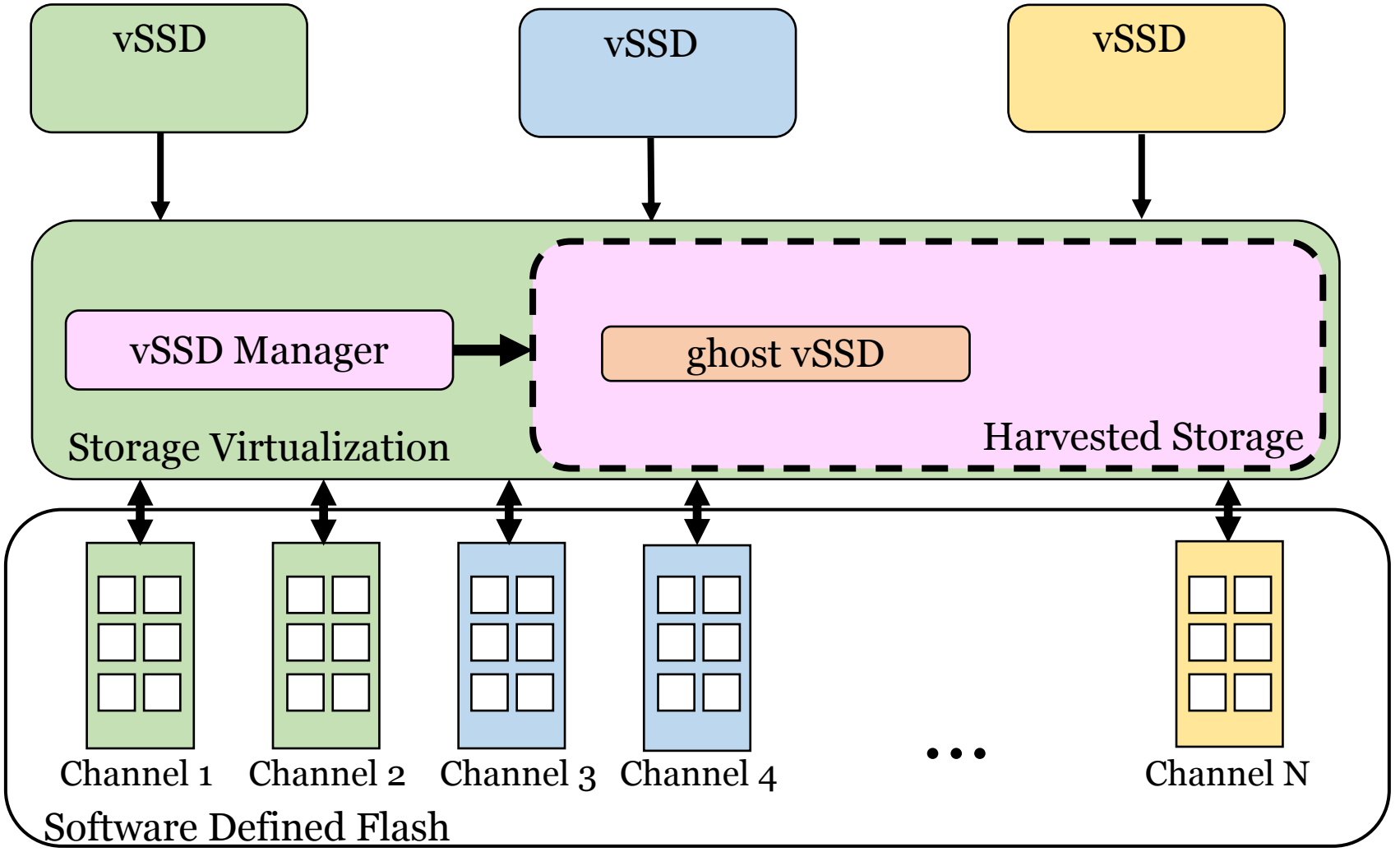


Hardware isolation between vSSDs

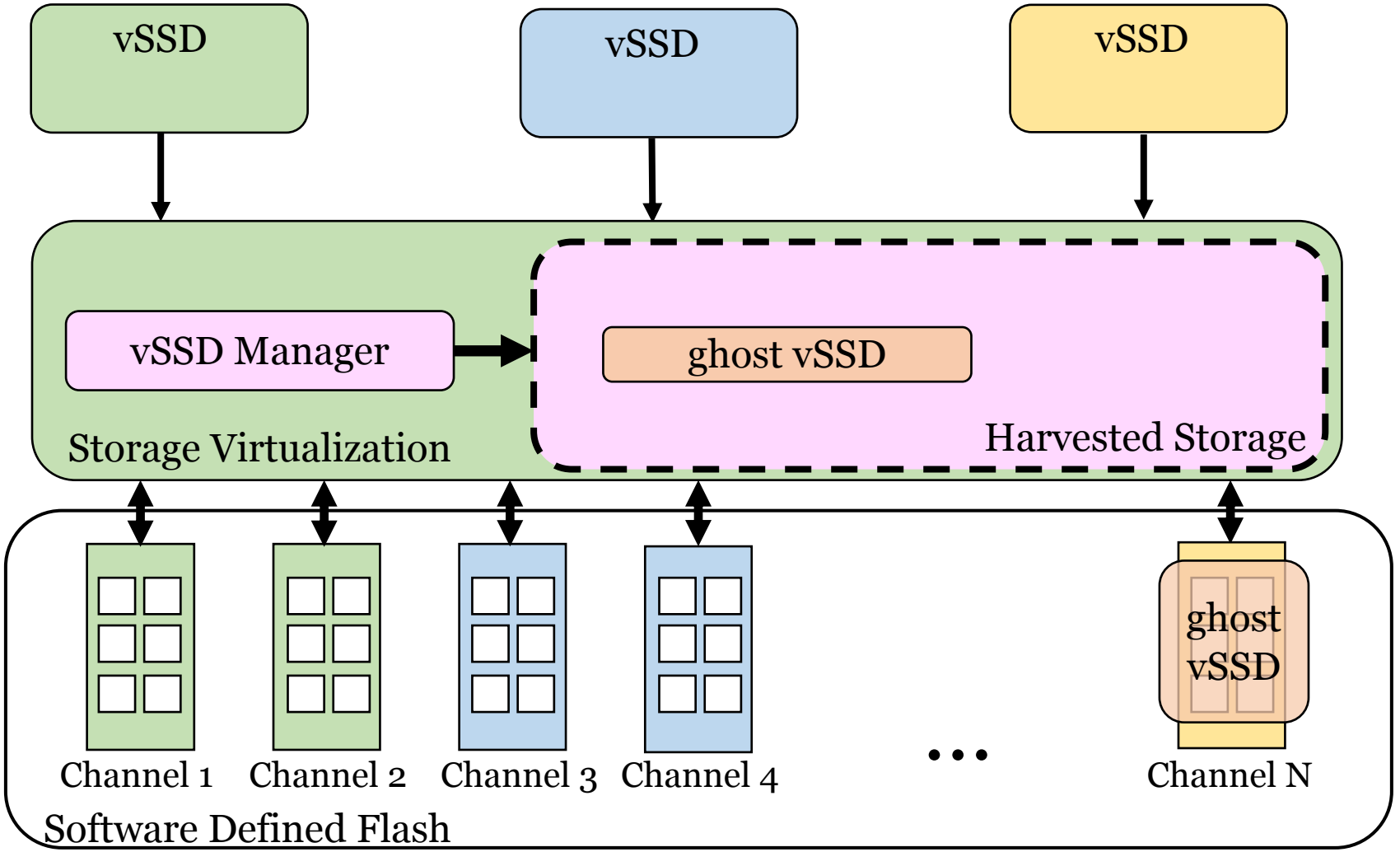
Ghost vSSD: A New vSSD Abstraction for Harvesting



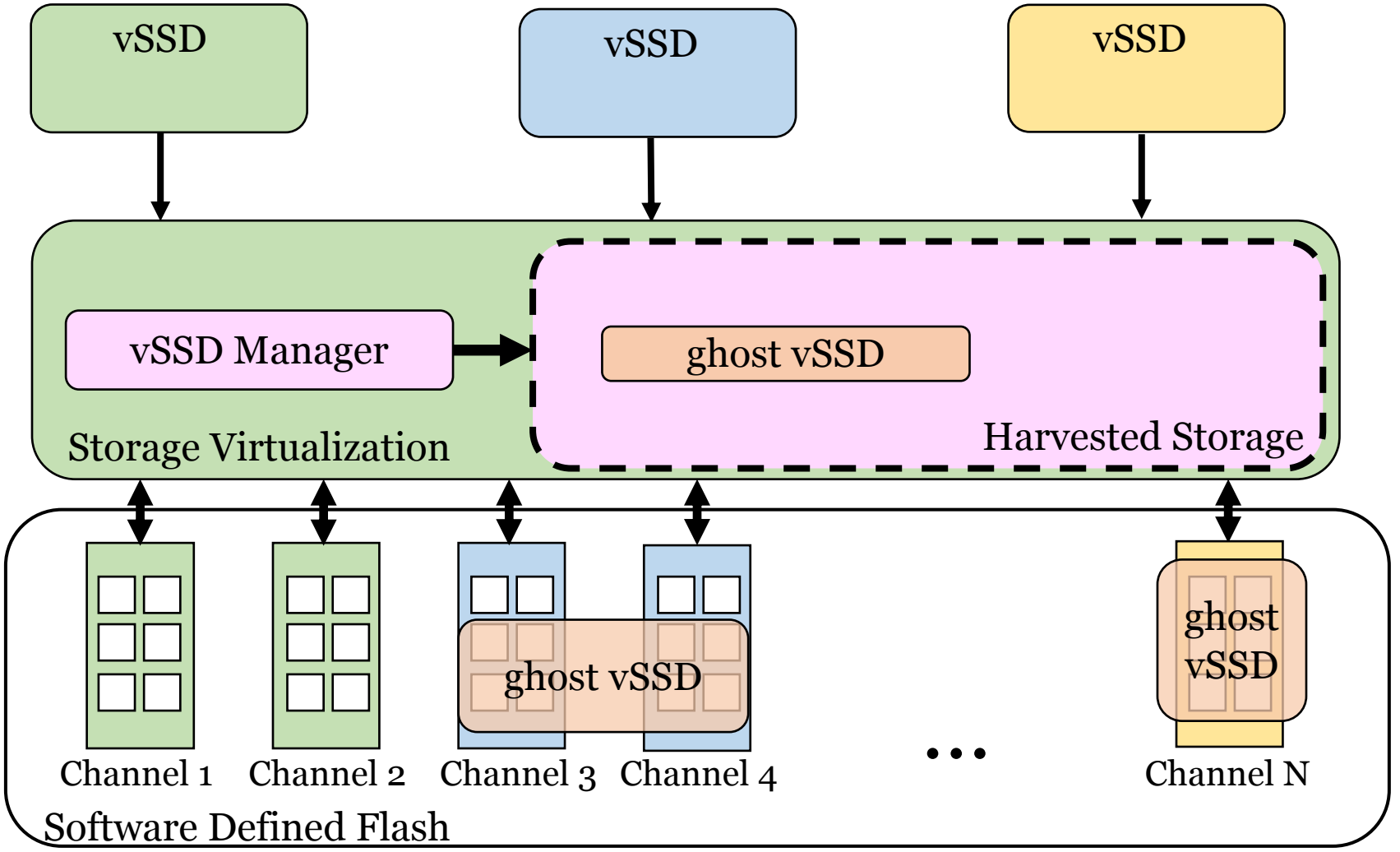
Ghost vSSD: A New vSSD Abstraction for Harvesting



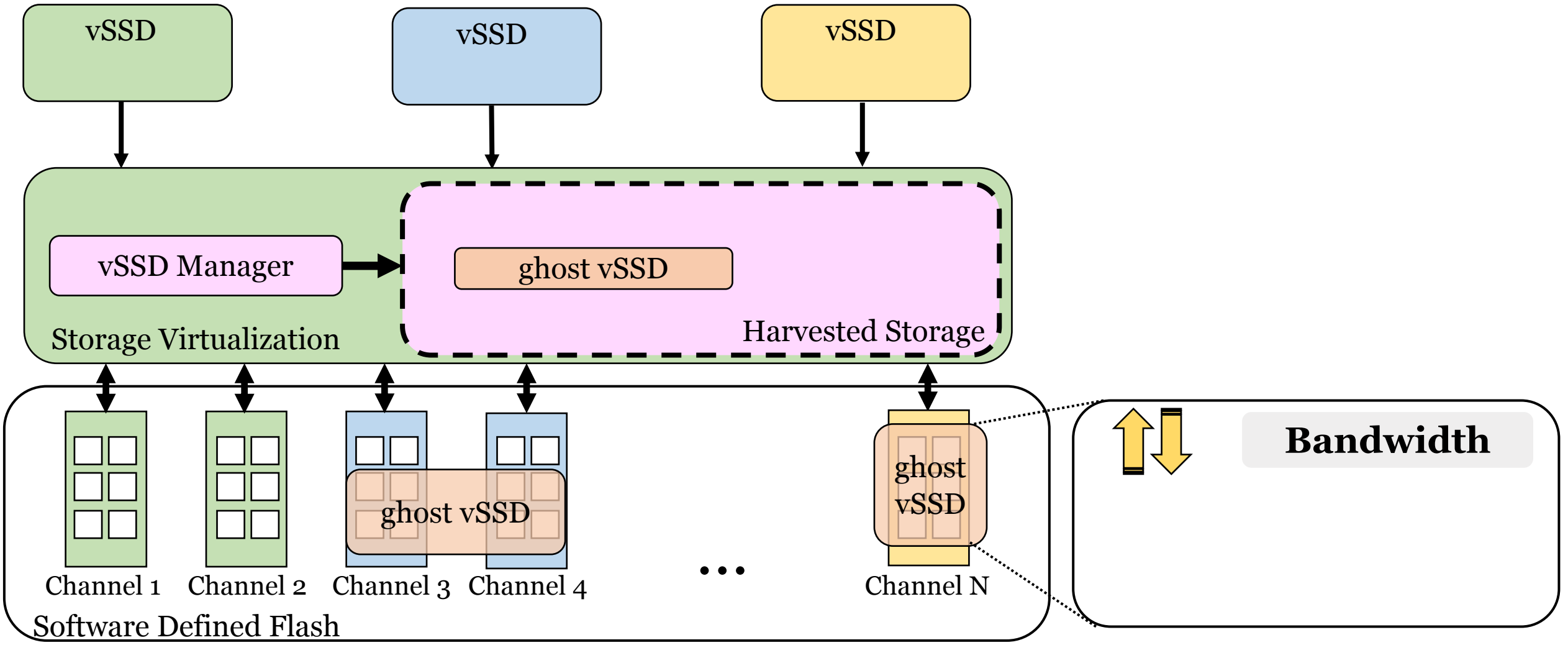
Ghost vSSD: A New vSSD Abstraction for Harvesting



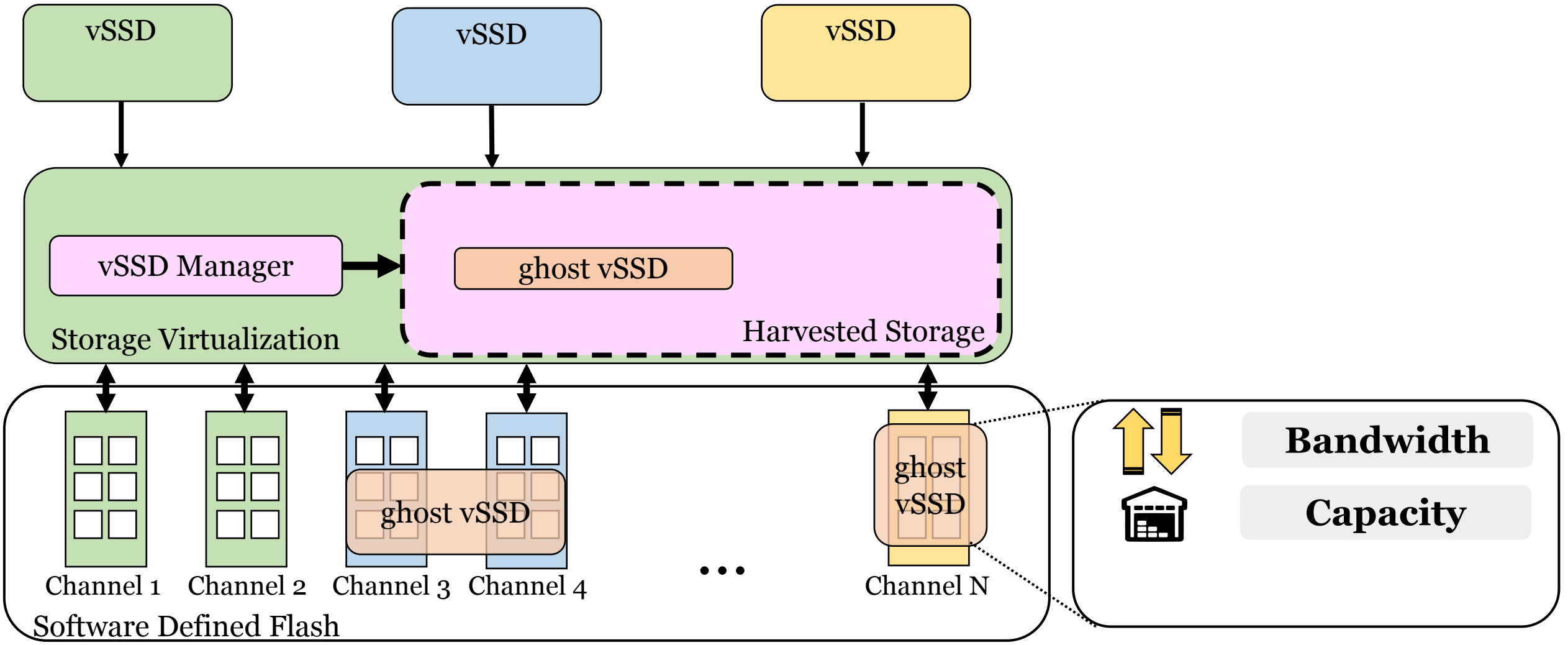
Ghost vSSD: A New vSSD Abstraction for Harvesting



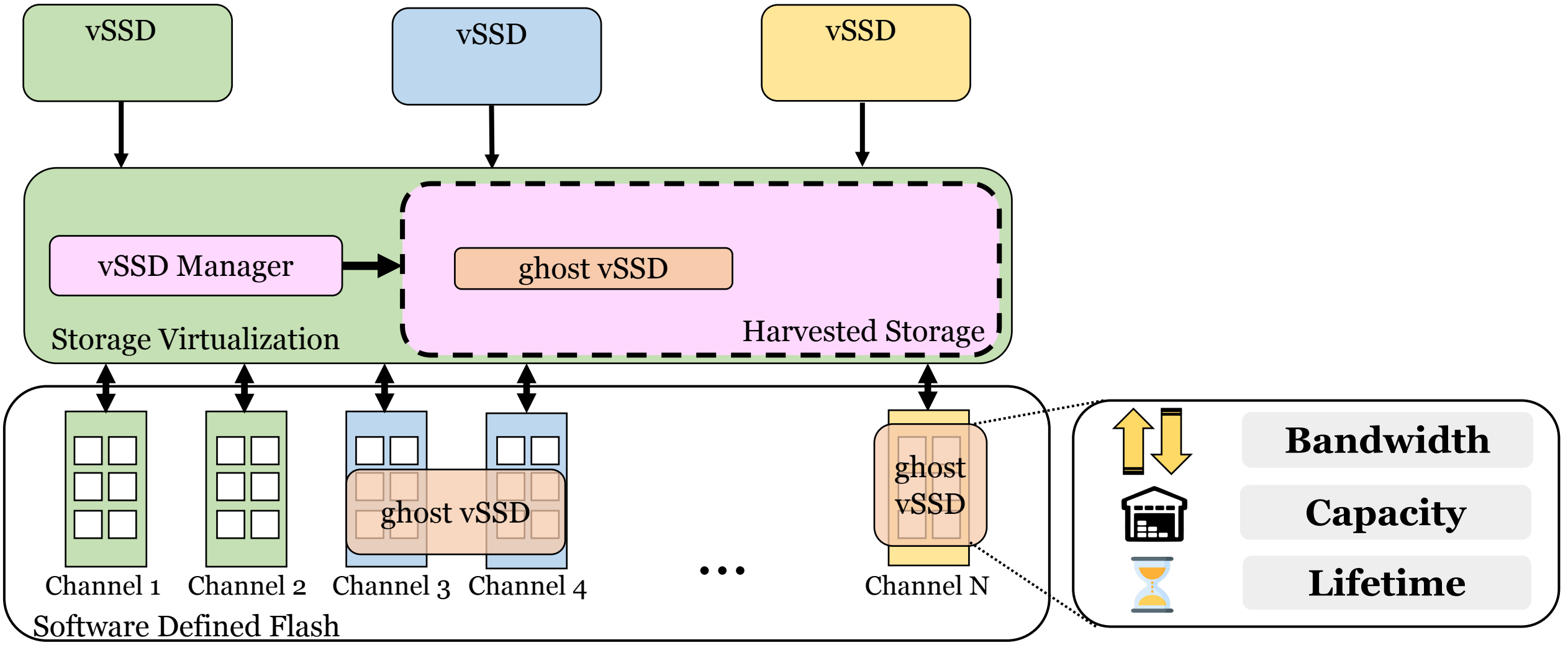
Ghost vSSD: A New vSSD Abstraction for Harvesting



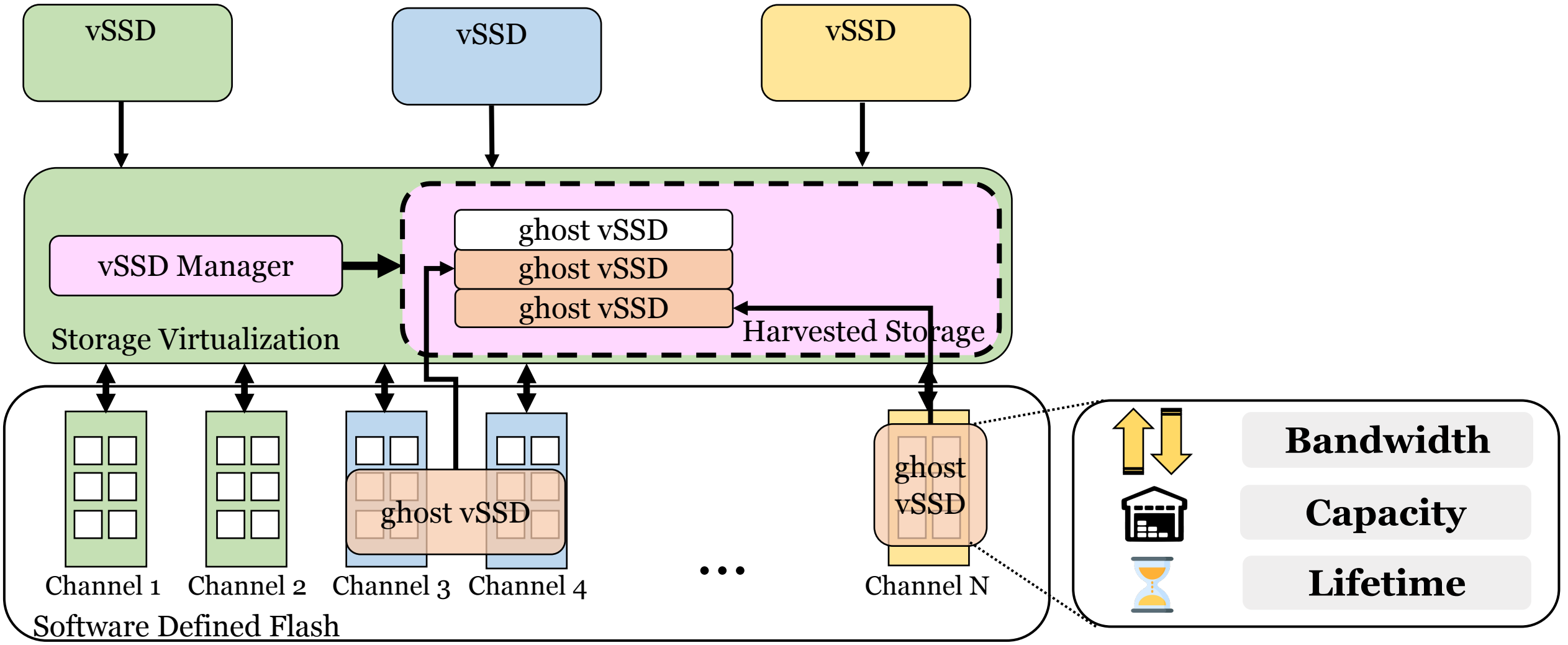
Ghost vSSD: A New vSSD Abstraction for Harvesting



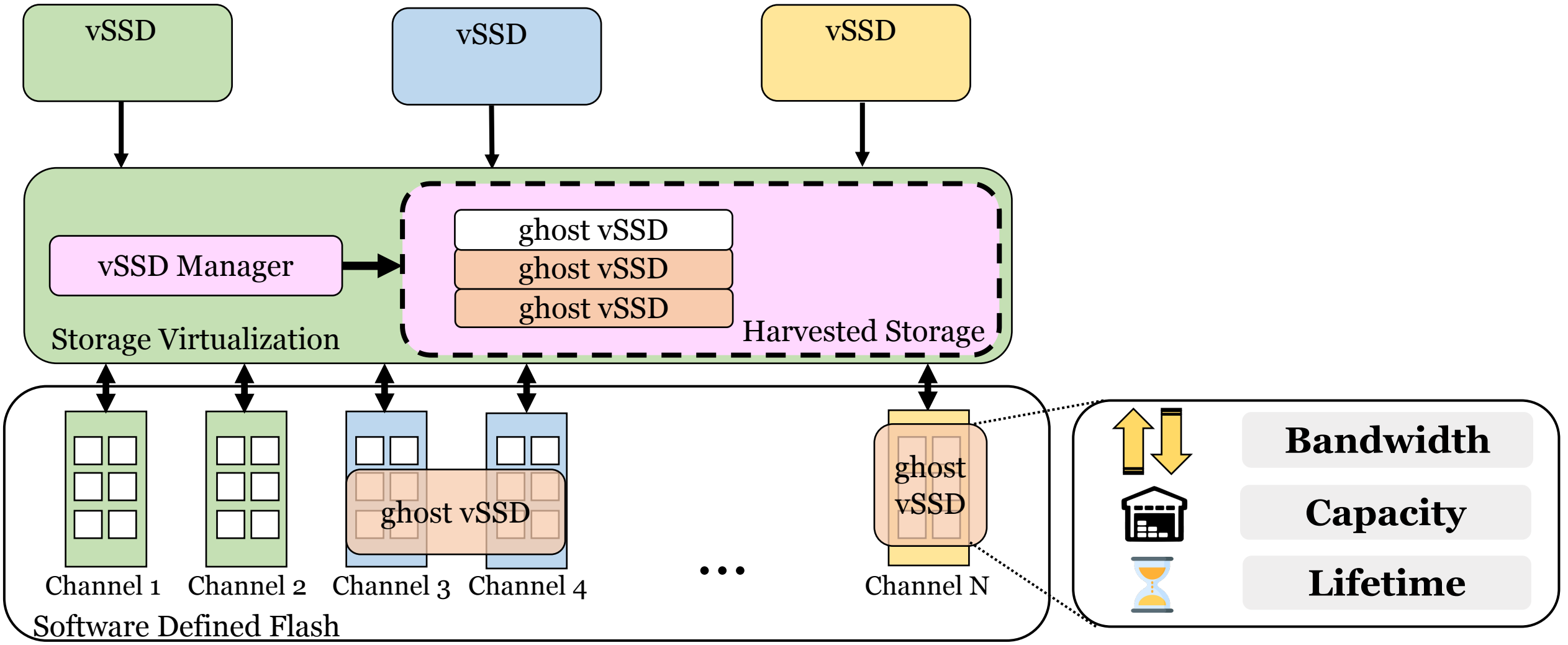
Ghost vSSD: A New vSSD Abstraction for Harvesting



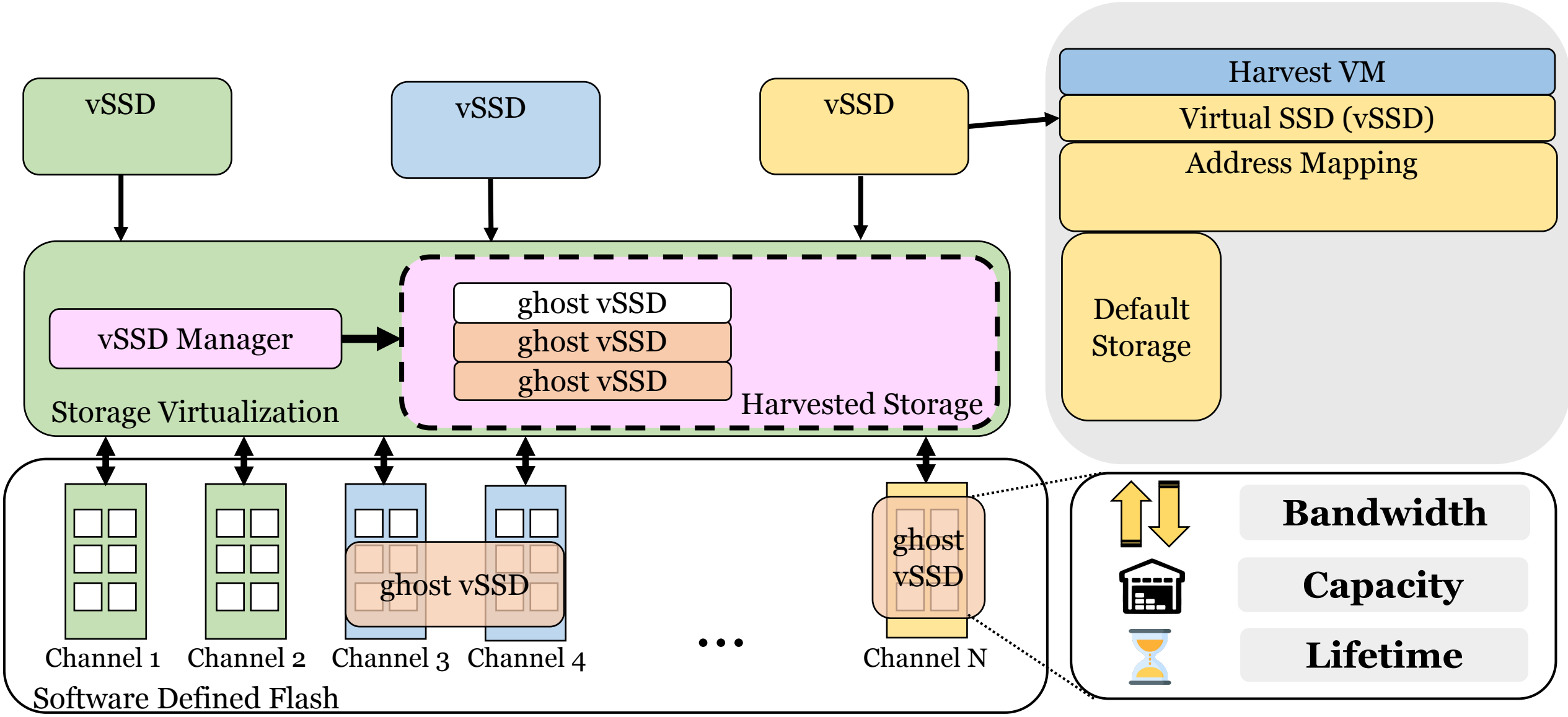
Ghost vSSD: A New vSSD Abstraction for Harvesting



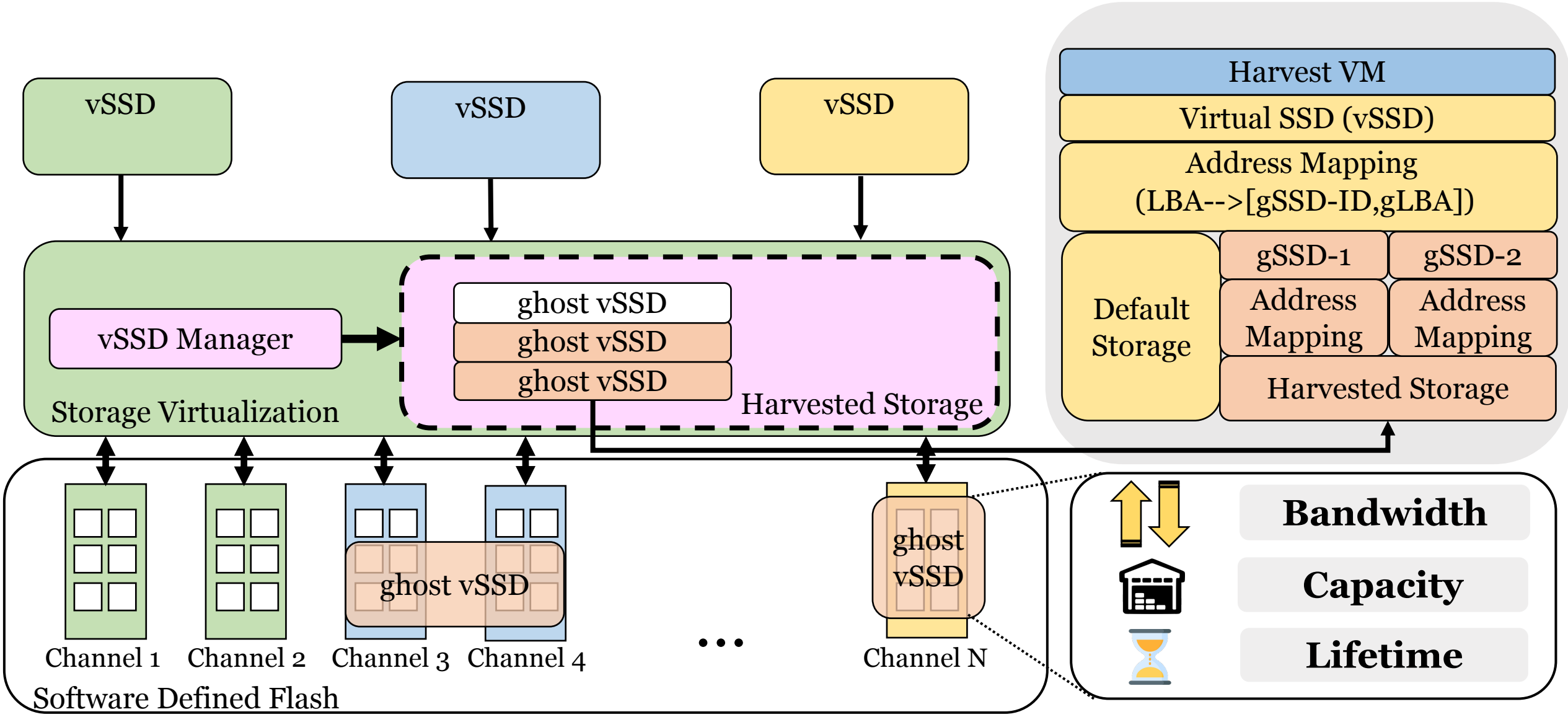
Ghost vSSD: A New vSSD Abstraction for Harvesting



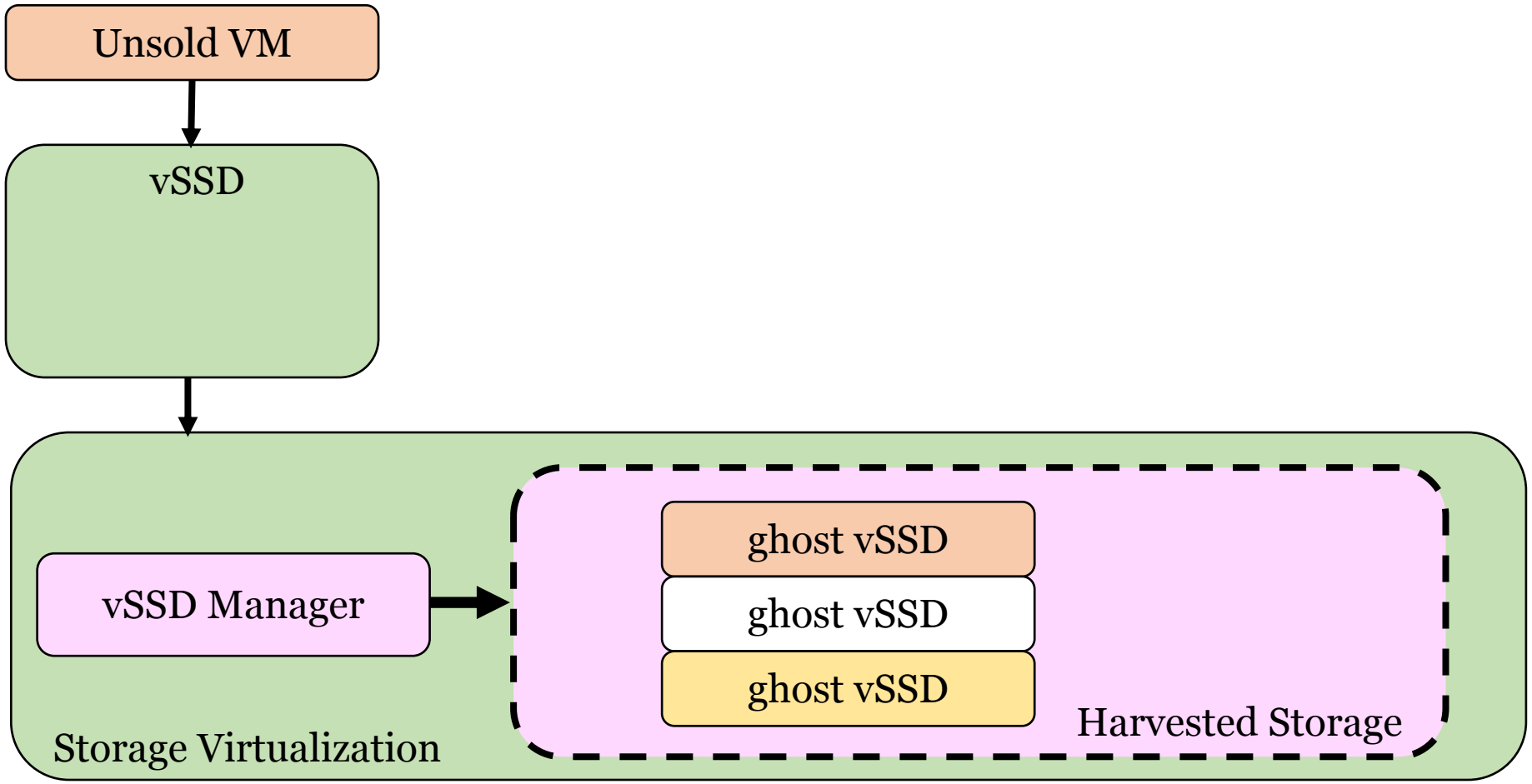
Ghost vSSD: A New vSSD Abstraction for Harvesting



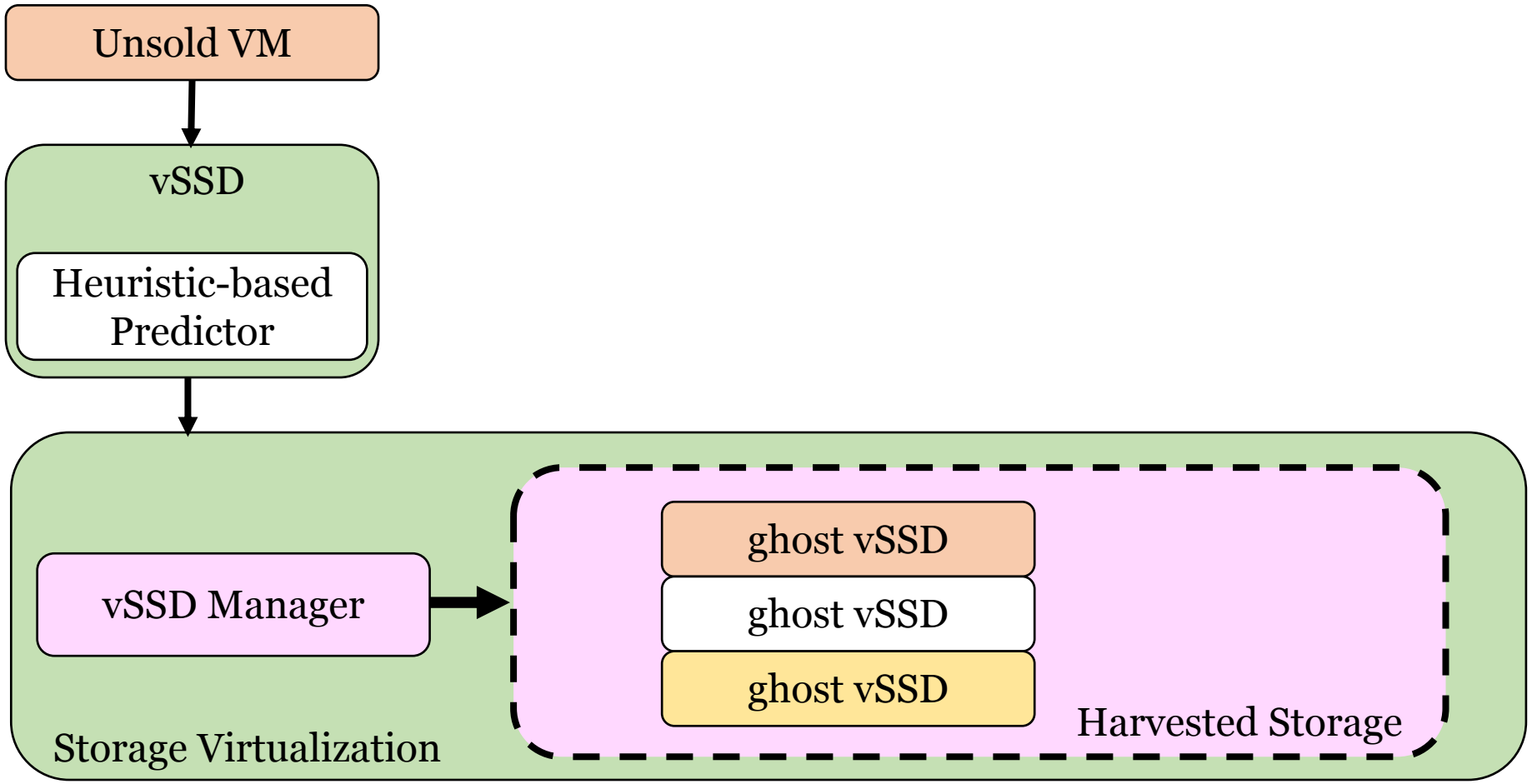
Ghost vSSD: A New vSSD Abstraction for Harvesting



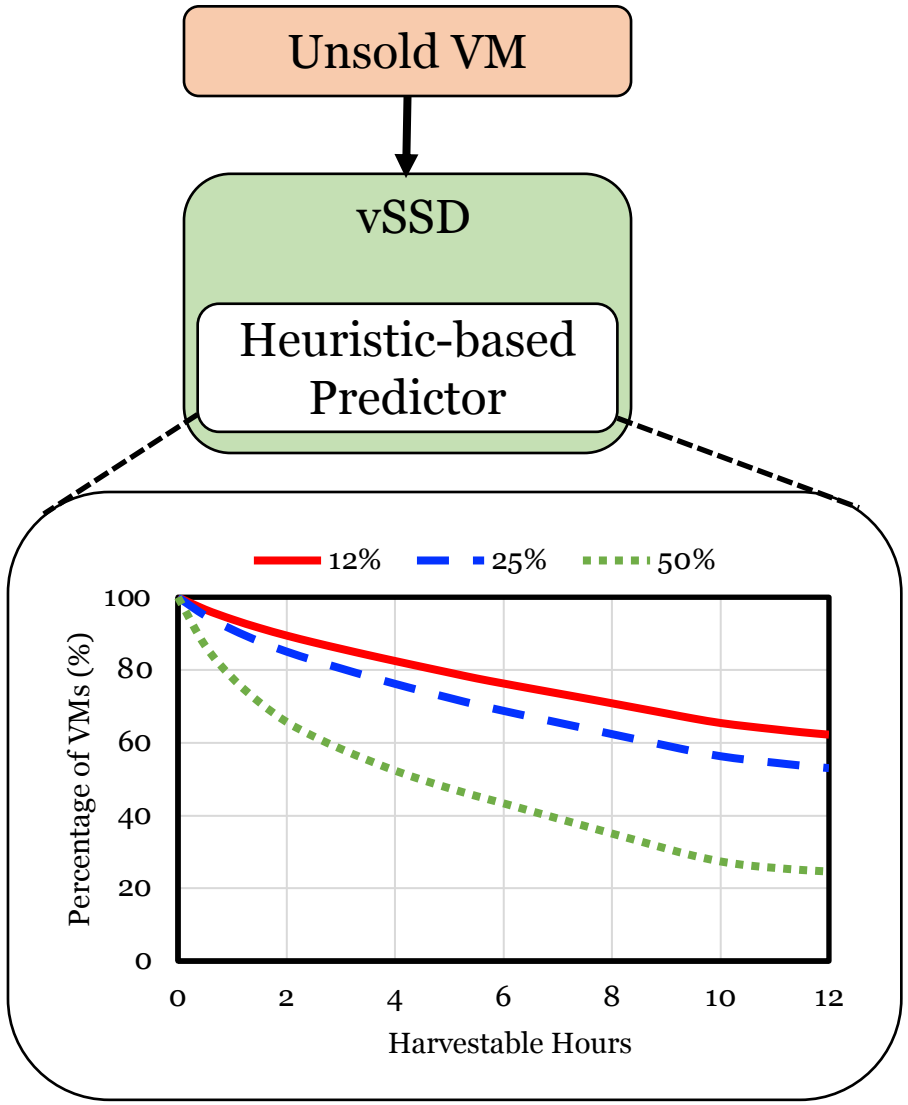
Facilitating Storage Harvesting with a Learning-Based Approach



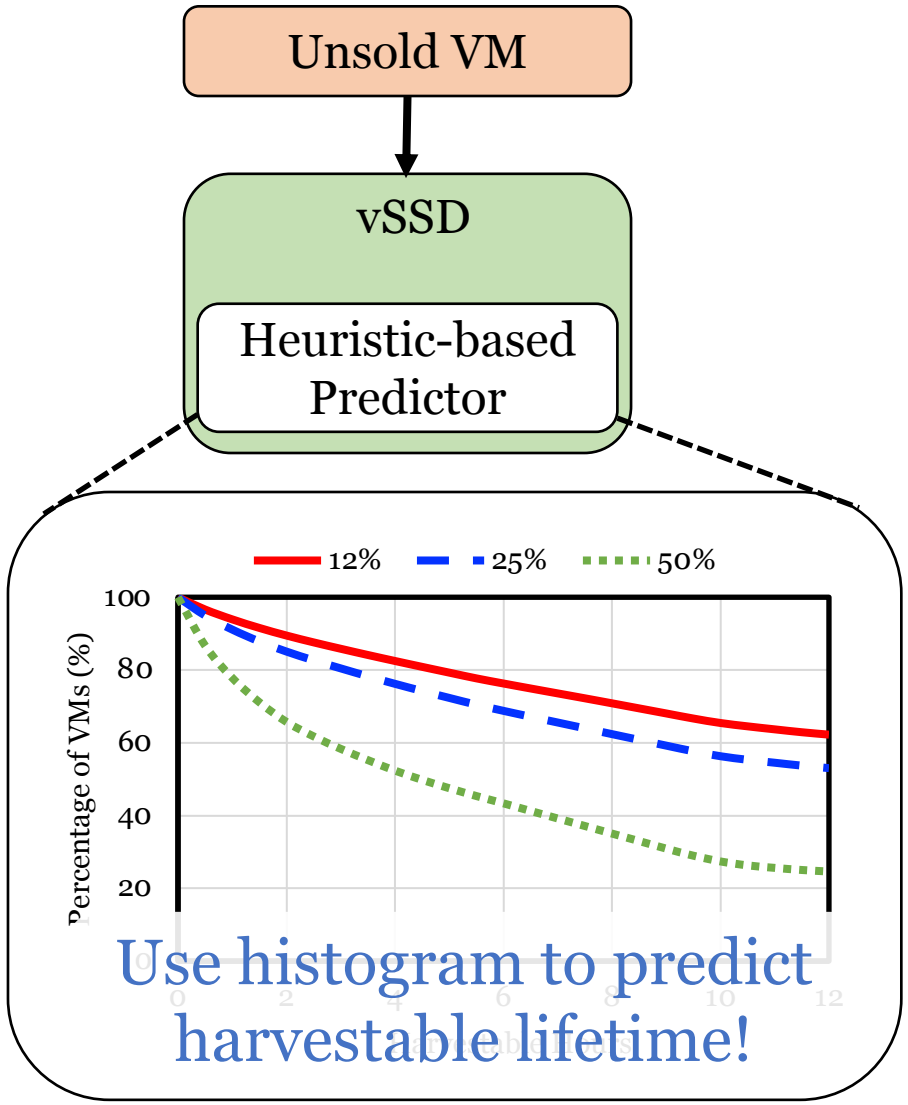
Facilitating Storage Harvesting with a Learning-Based Approach



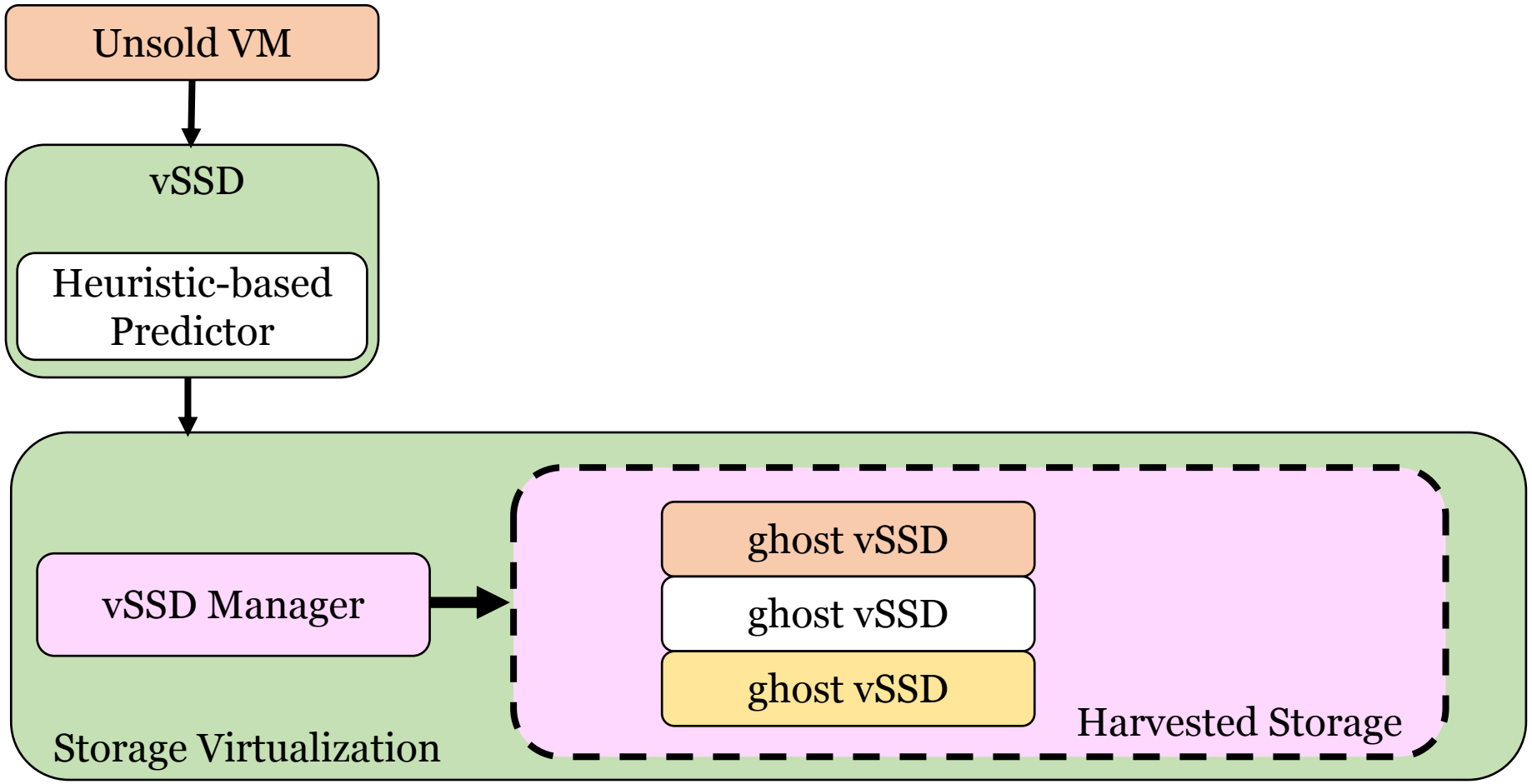
Facilitating Storage Harvesting with a Learning-Based Approach



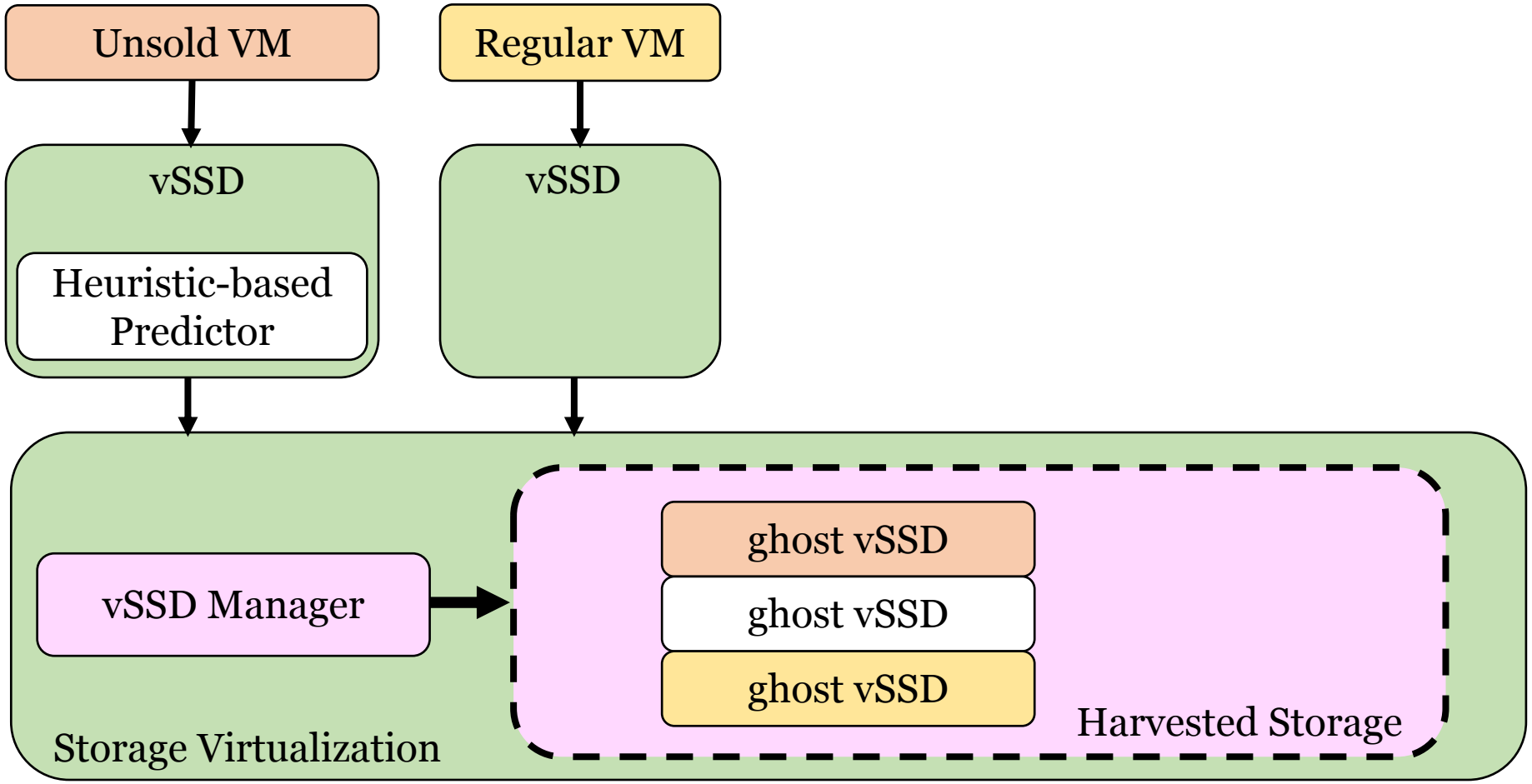
Facilitating Storage Harvesting with a Learning-Based Approach



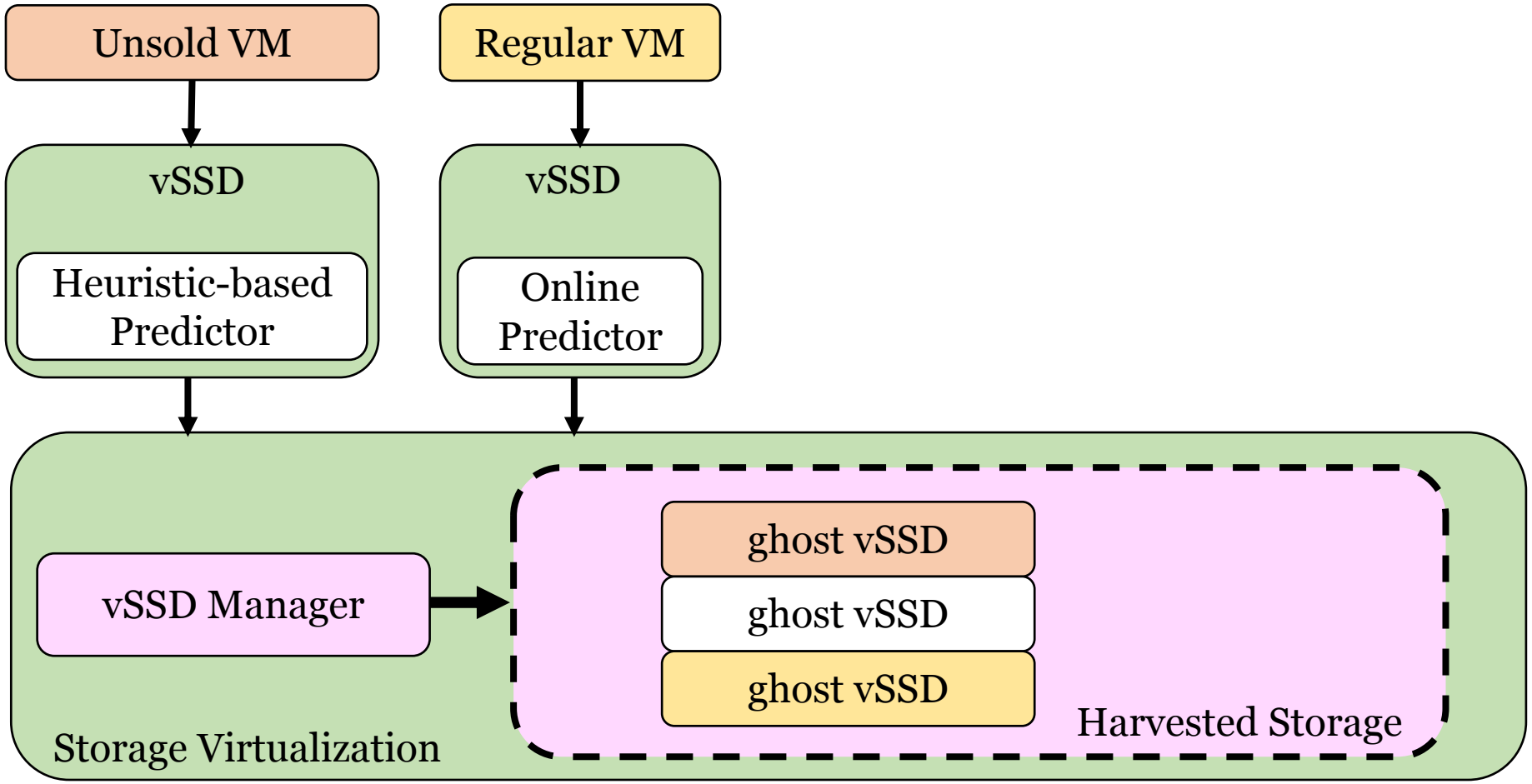
Facilitating Storage Harvesting with a Learning-Based Approach



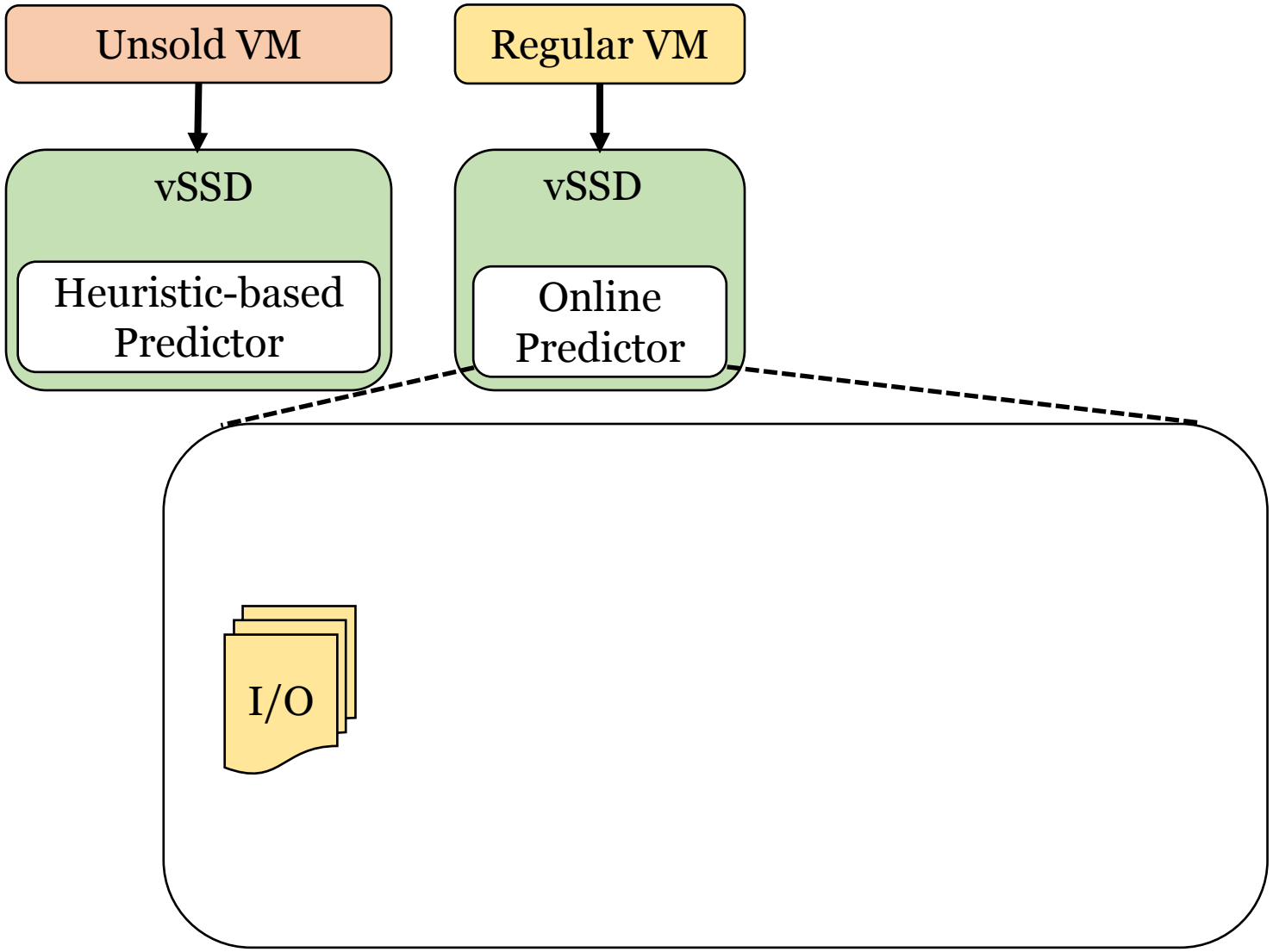
Facilitating Storage Harvesting with a Learning-Based Approach



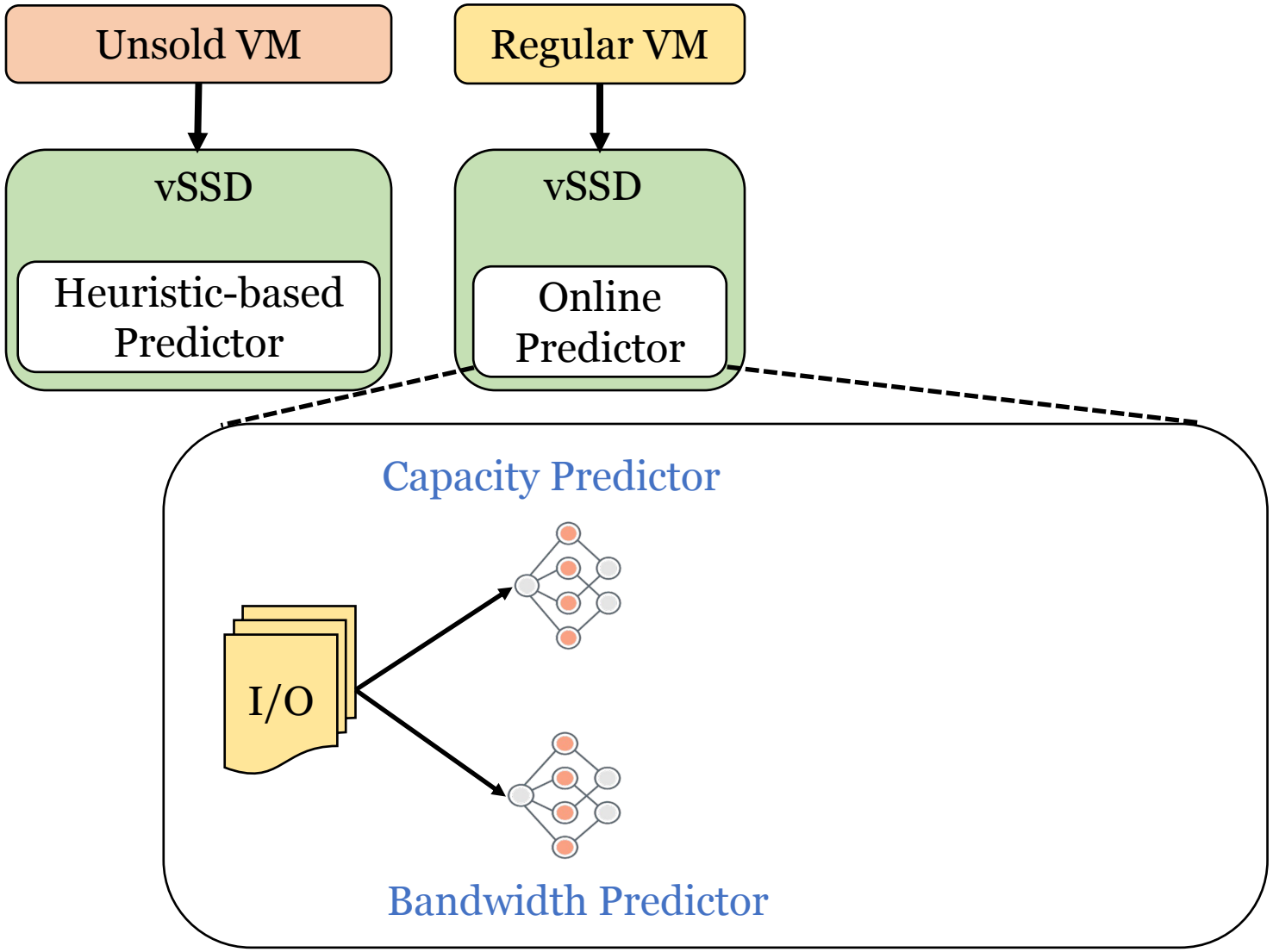
Facilitating Storage Harvesting with a Learning-Based Approach



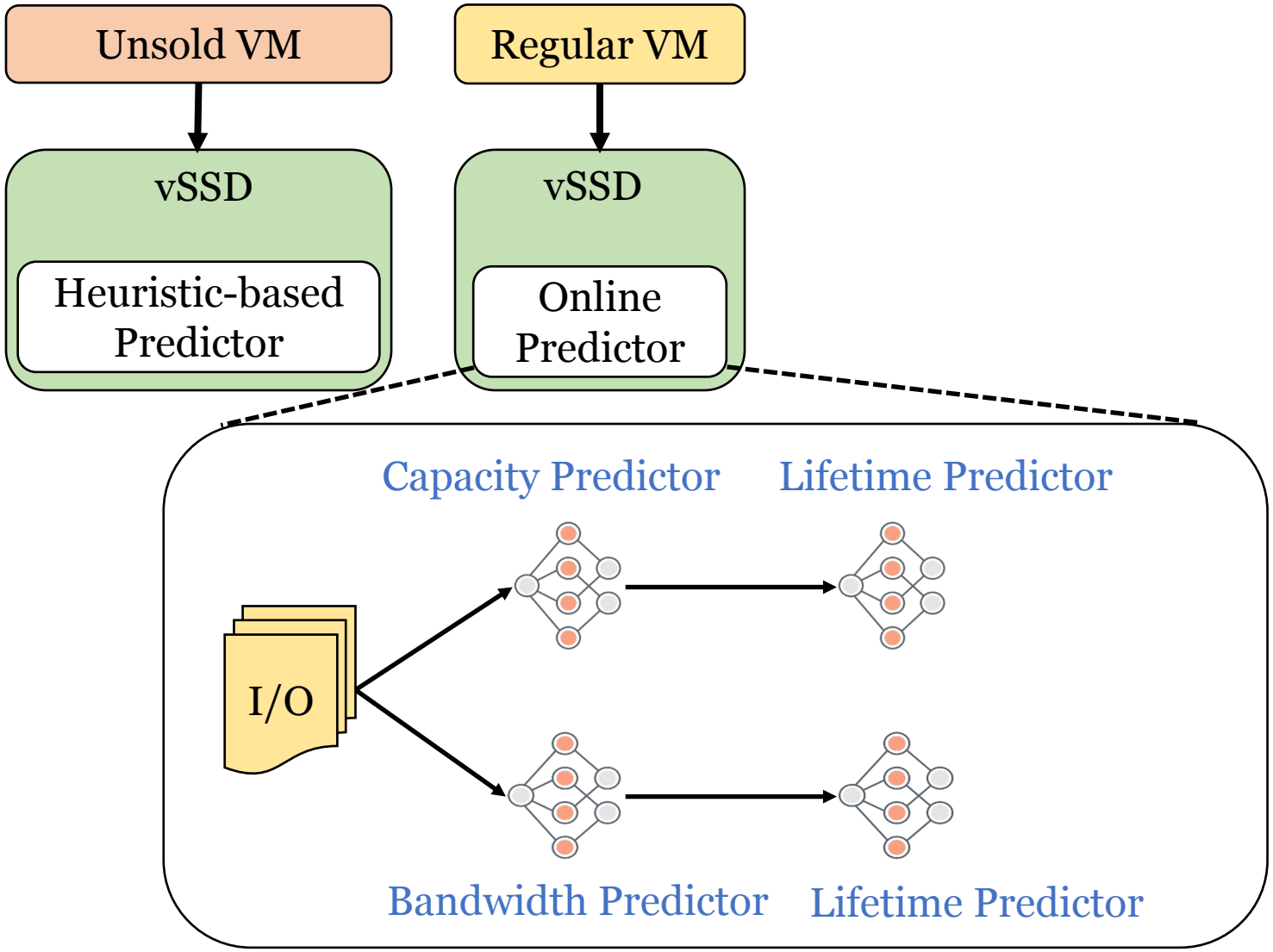
Facilitating Storage Harvesting with a Learning-Based Approach



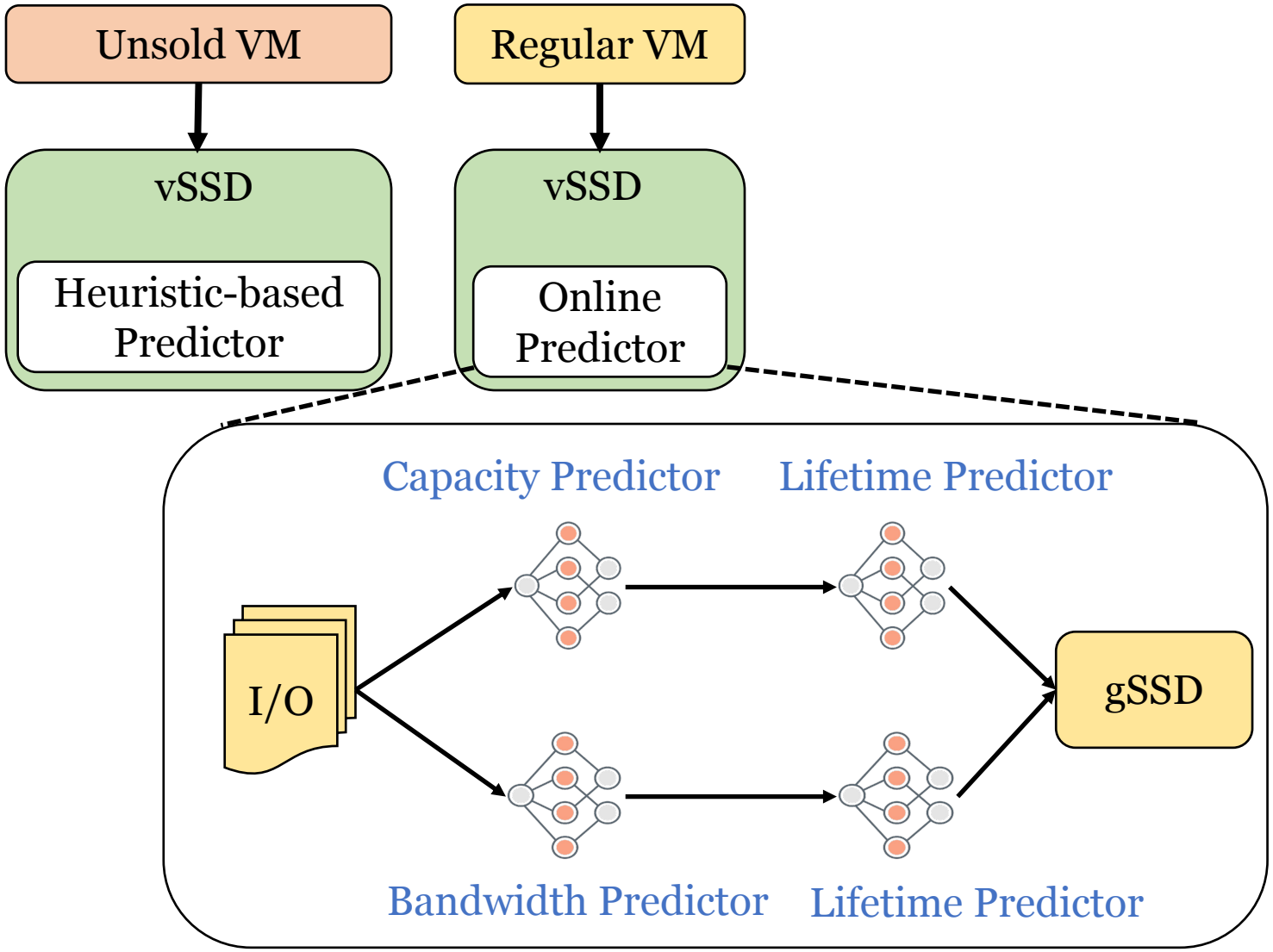
Facilitating Storage Harvesting with a Learning-Based Approach



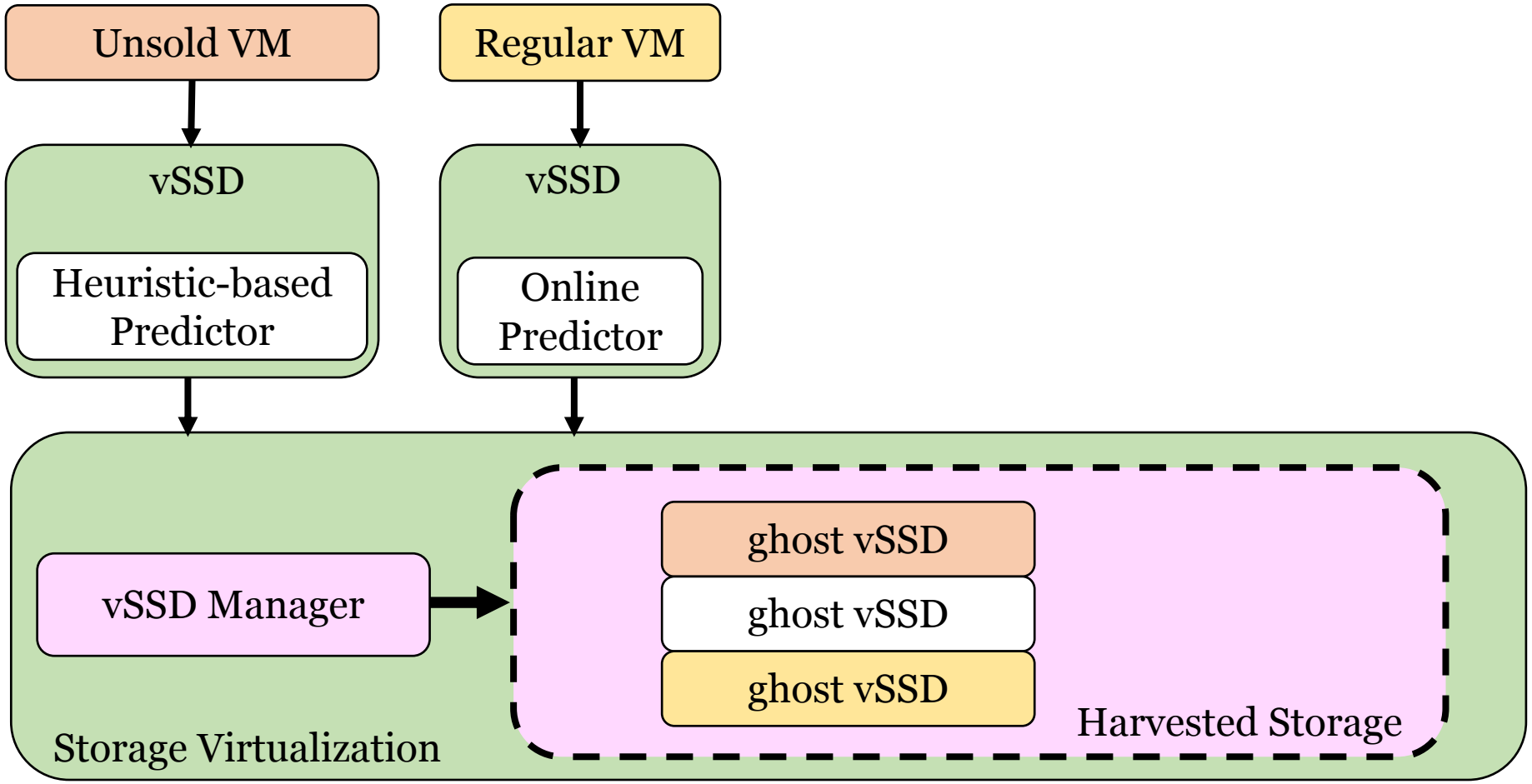
Facilitating Storage Harvesting with a Learning-Based Approach



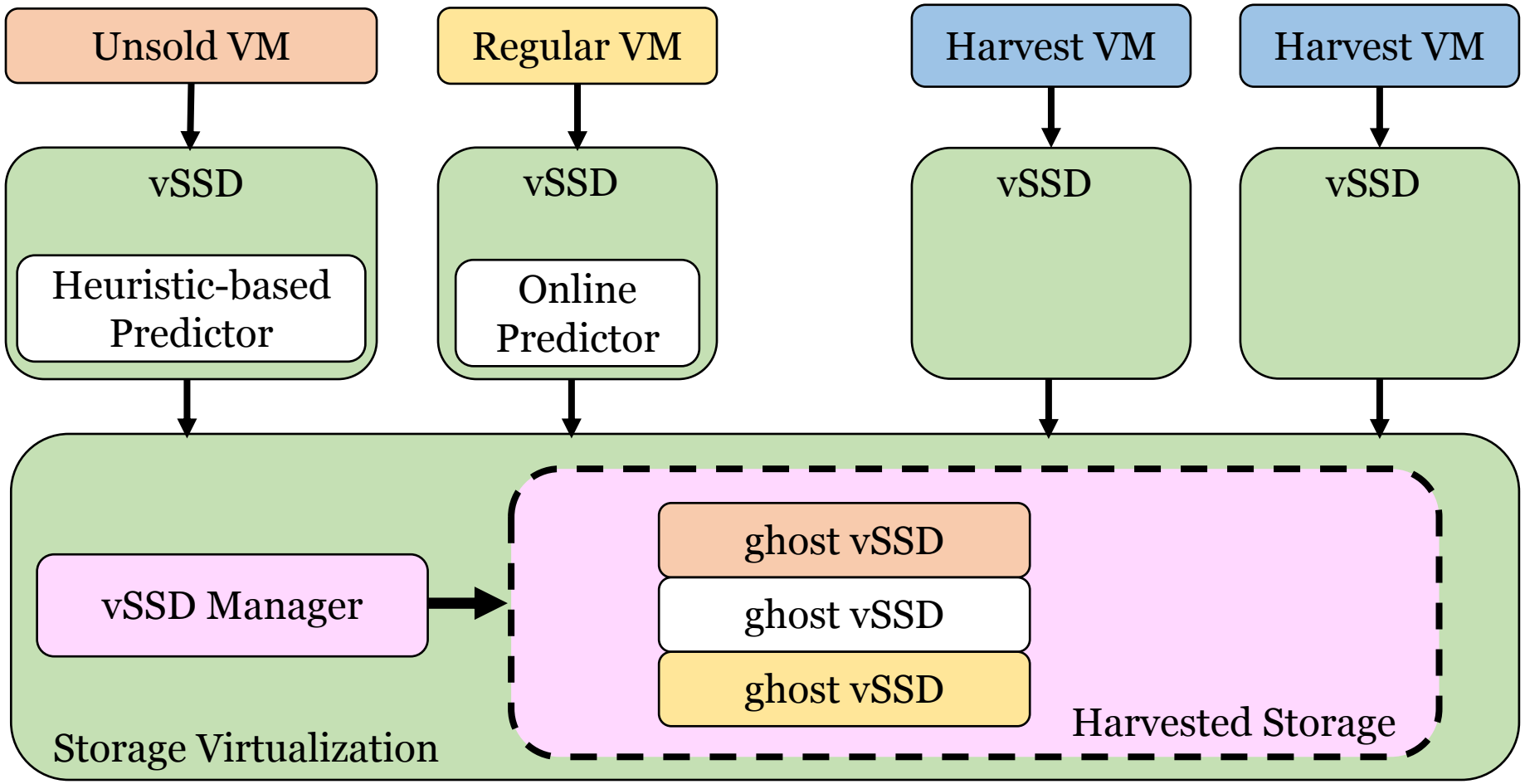
Facilitating Storage Harvesting with a Learning-Based Approach



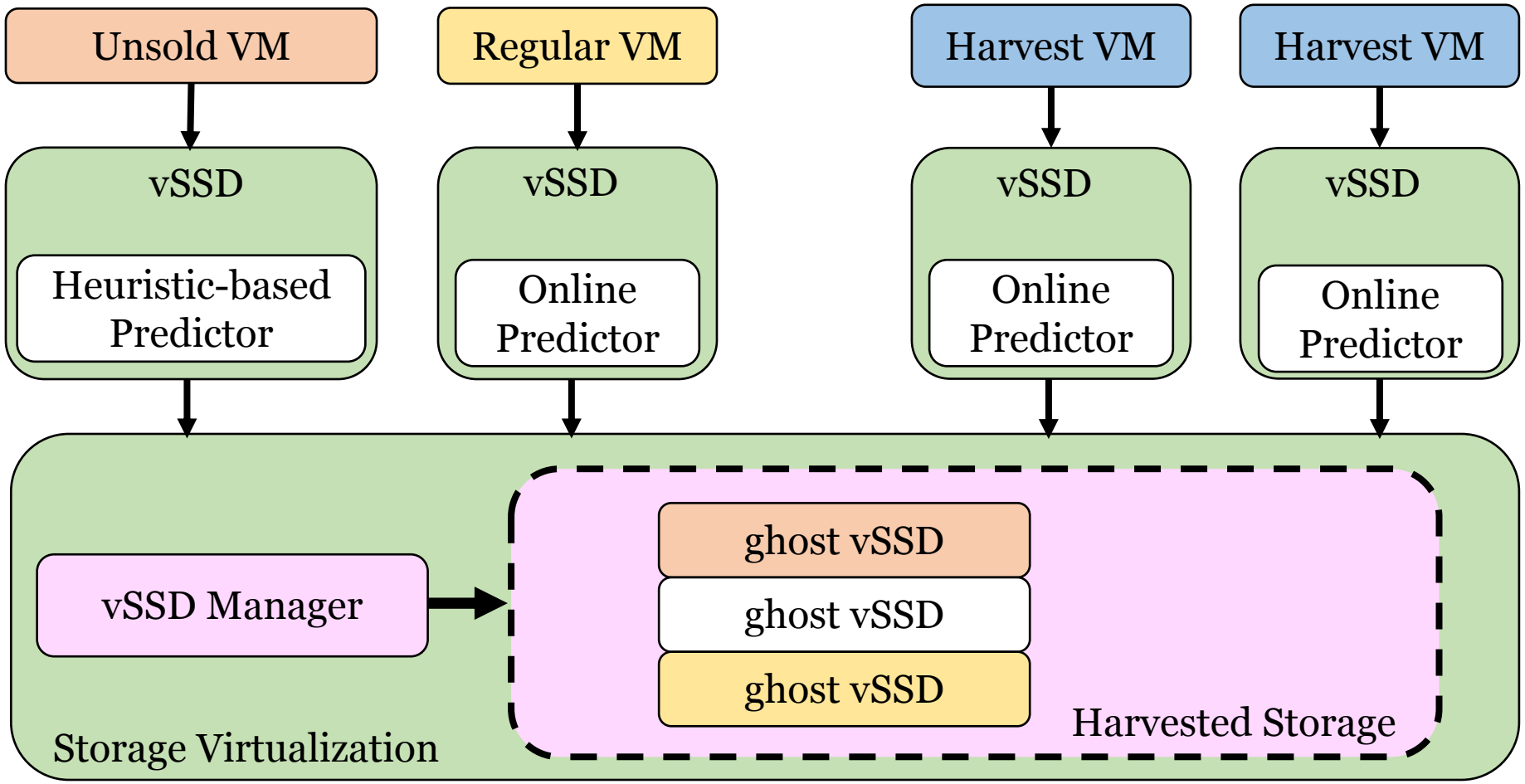
Facilitating Storage Harvesting with a Learning-Based Approach



Facilitating Storage Harvesting with a Learning-Based Approach



Facilitating Storage Harvesting with a Learning-Based Approach



Handling Mispredictions in BlockFlex

Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	

Handling Mispredictions in BlockFlex

Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	

Handling Mispredictions in BlockFlex

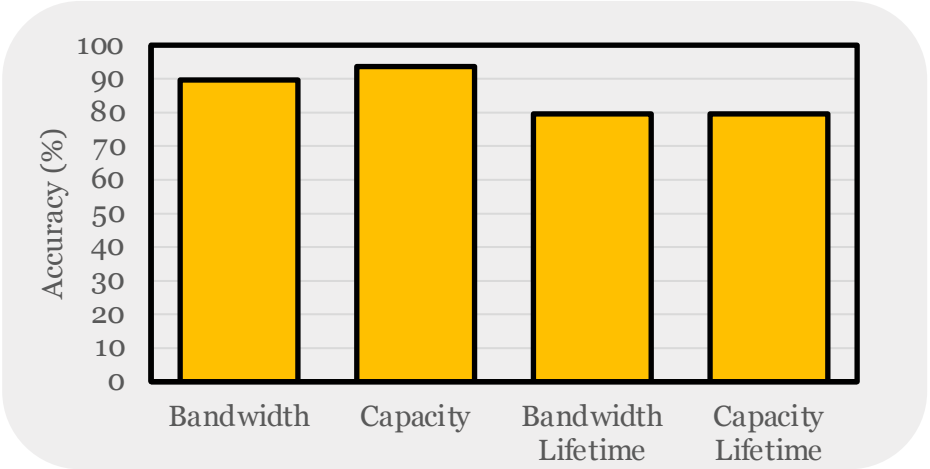
Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	

Handling Mispredictions in BlockFlex

Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	

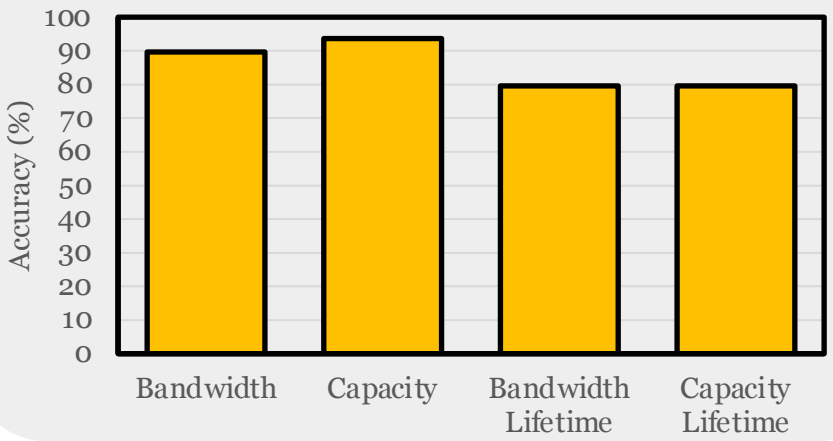
Handling Mispredictions in BlockFlex

Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	

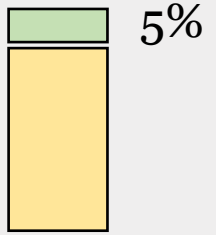


Handling Mispredictions in BlockFlex

Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	



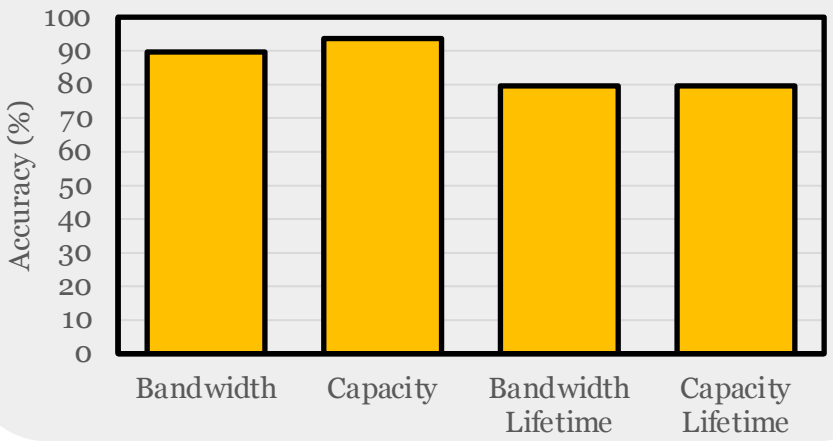
+



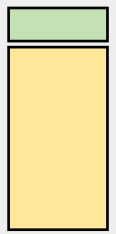
Add over-provisioning to avoid reclamations

Handling Mispredictions in BlockFlex

Harvestable Storage	Demanded Storage	Wasted Resources	Under Harvesting	Early Reclamation
Over-Prediction	Over-Prediction	✗		✗
Over-Prediction	Under-Prediction		✗	✗
Under-Prediction	Over-Prediction	✗		
Under-Prediction	Under-Prediction	✗	✗	



+



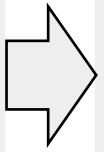
Add over-provisioning to avoid reclamations

=

93-96% Accuracy!

BlockFlex: Enabling Storage Harvesting in Modern Cloud

Insufficient Virtualization Support



Extend Storage Virtualization with **Ghost vSSD Abstraction**

Data Security Concern

Enable Fine-Grained **Hardware Isolation** and **Block Erasing**

Storage Performance Interference

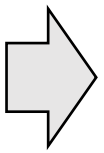
Minimize the **Impact of Harvesting** on Regular VMs

BlockFlex: Enabling Storage Harvesting in Modern Cloud

Insufficient Virtualization Support

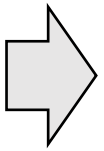
Extend Storage Virtualization with **Ghost vSSD Abstraction**

Data Security Concern



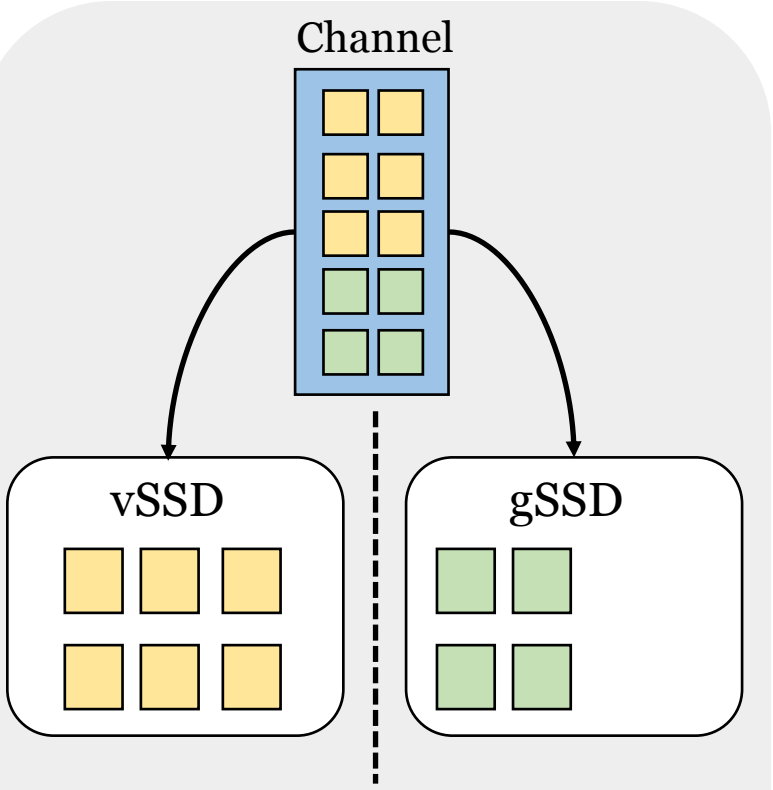
Enable Fine-Grained **Hardware Isolation** and **Block Erasing**

Storage Performance Interference



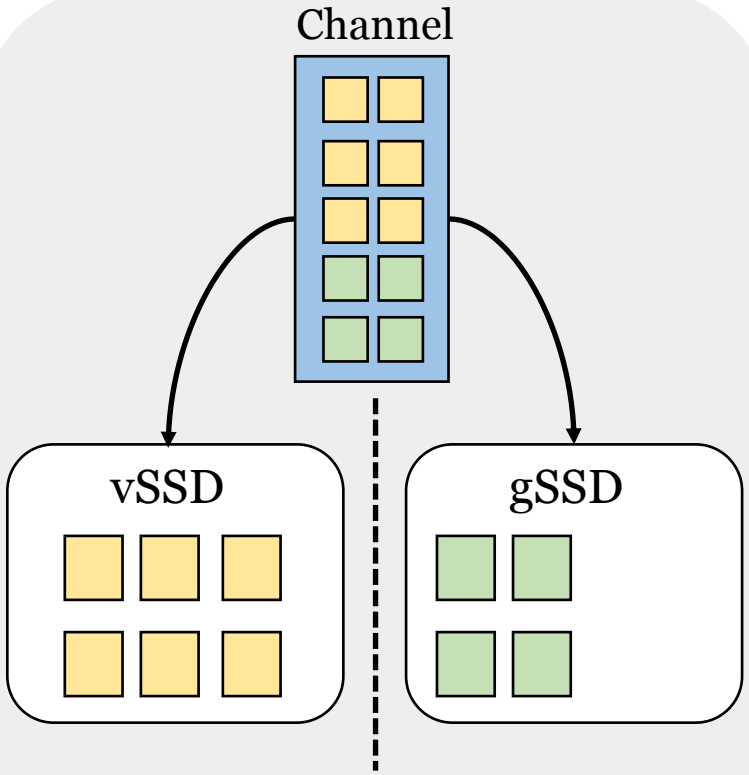
Minimize the **Impact of Harvesting** on Regular VMs

Enforcing Data Security in BlockFlex

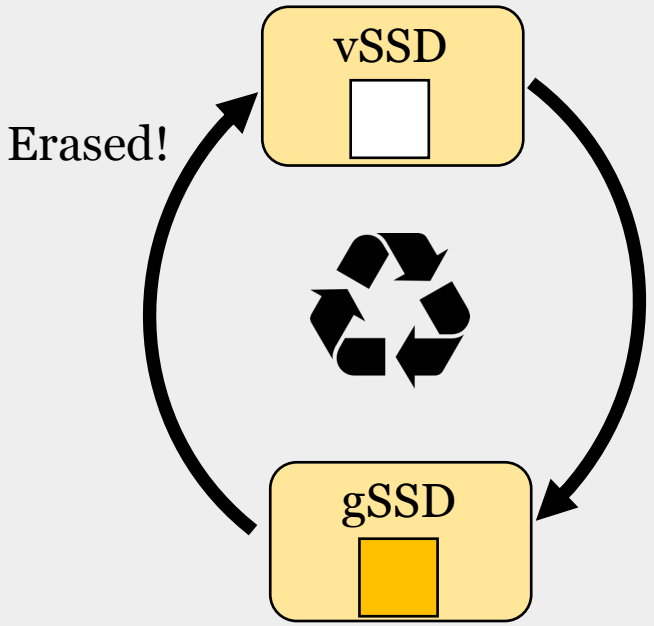


Each gSSD is **hardware isolated** from its vSSD

Enforcing Data Security in BlockFlex

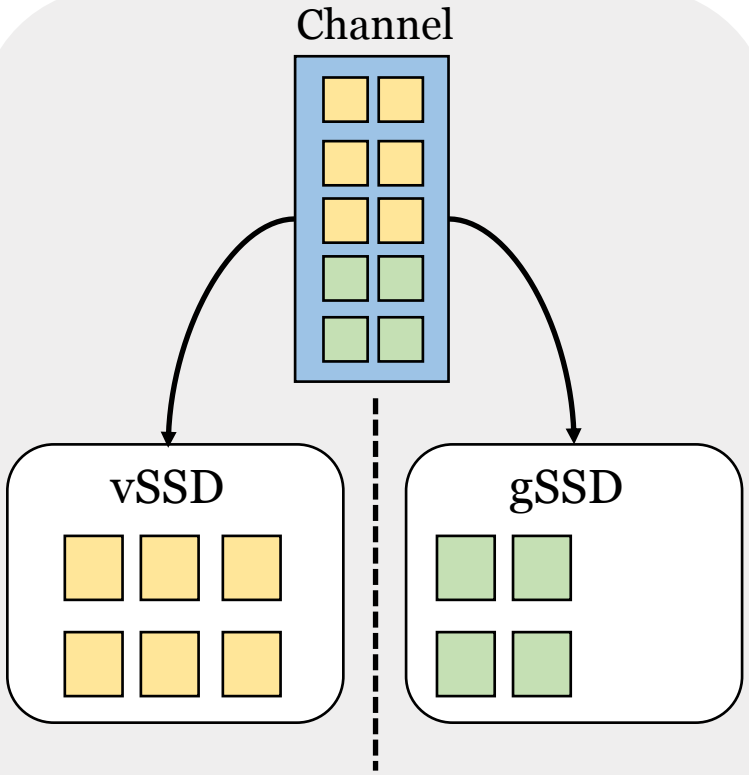


Each gSSD is **hardware isolated** from its vSSD

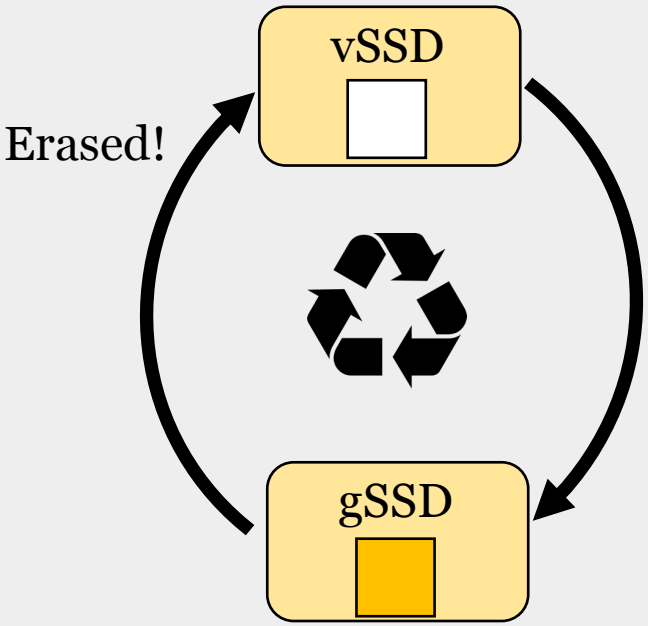


All blocks are **erased** after harvesting

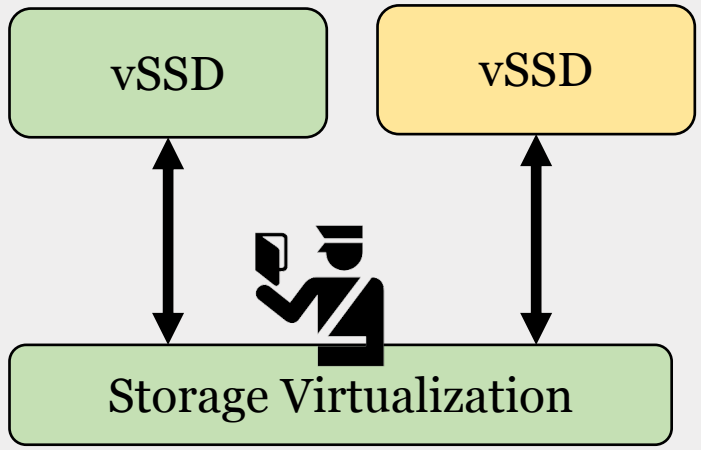
Enforcing Data Security in BlockFlex



Each gSSD is **hardware isolated** from its vSSD

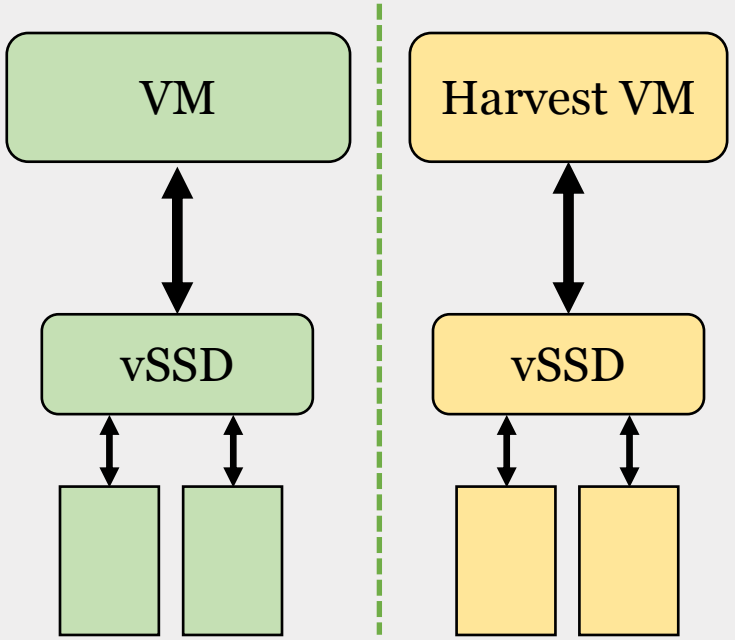


All blocks are **erased** after harvesting



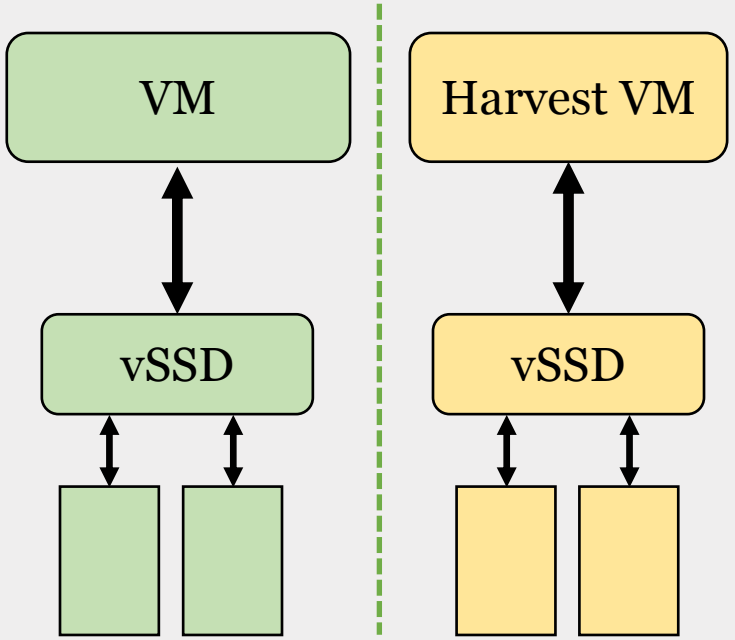
Permission checks within storage virtualization

Guaranteed Performance Isolation in BlockFlex

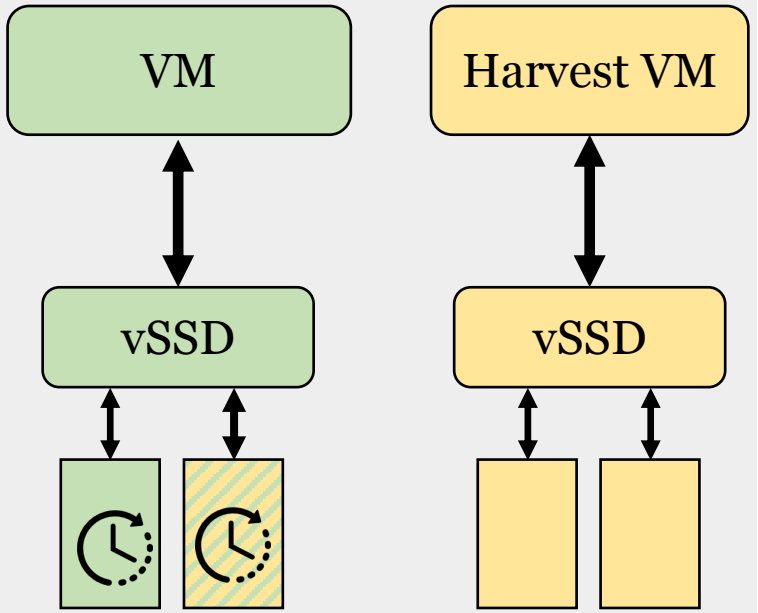


vSSDs are mapped to **independent** channels

Guaranteed Performance Isolation in BlockFlex

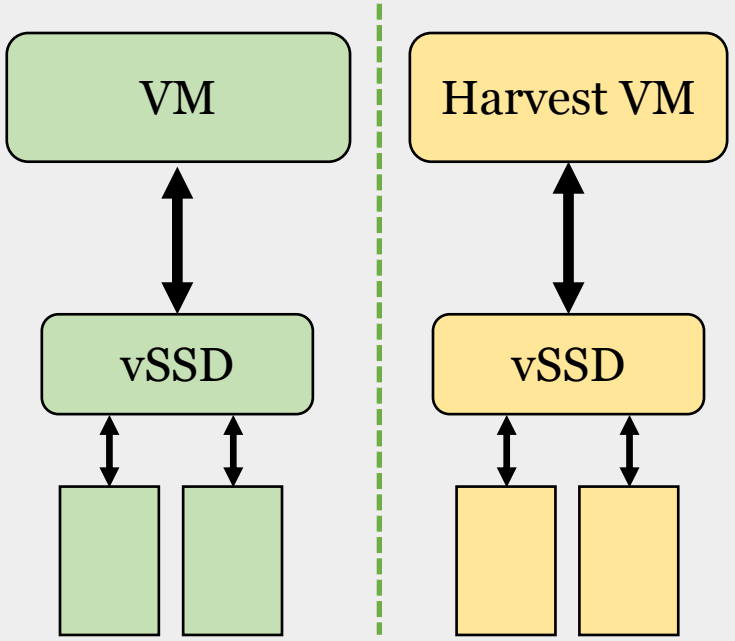


vSSDs are mapped to **independent** channels

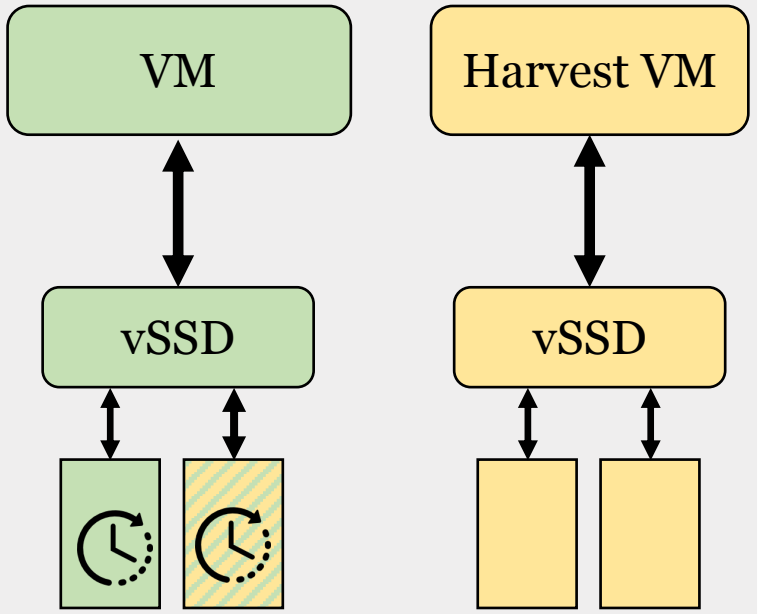


Only harvest **idle resources**

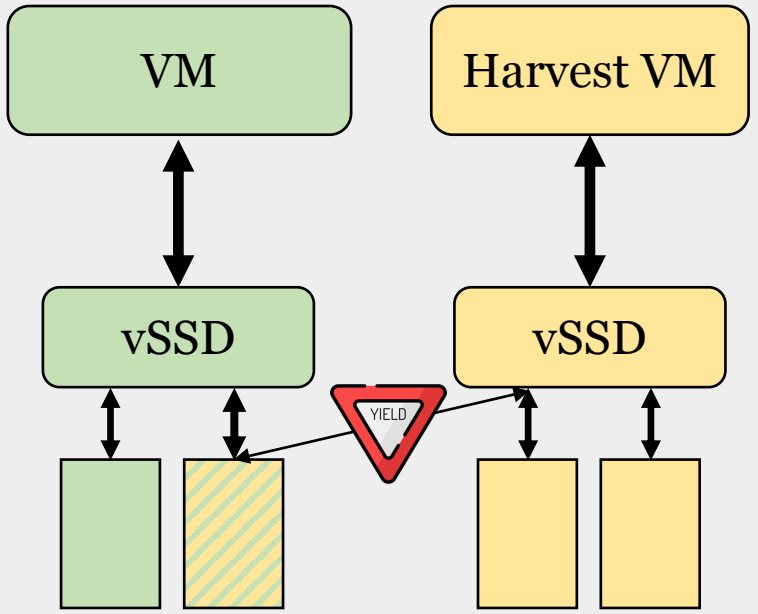
Guaranteed Performance Isolation in BlockFlex



vSSDs are mapped to **independent** channels



Only harvest **idle resources**



Prioritized I/O for regular VMs

BlockFlex Implementation

Programmable SSD

1 TB
16 Channels
16 KB/page
70 MB/s per channel

BlockFlex Implementation

Programmable SSD

1 TB
16 Channels
16 KB/page
70 MB/s per channel

Regular VM Workloads

YCSB (Workloads A-E)

BlockFlex Implementation

Programmable SSD

1 TB
16 Channels
16 KB/page
70 MB/s per channel

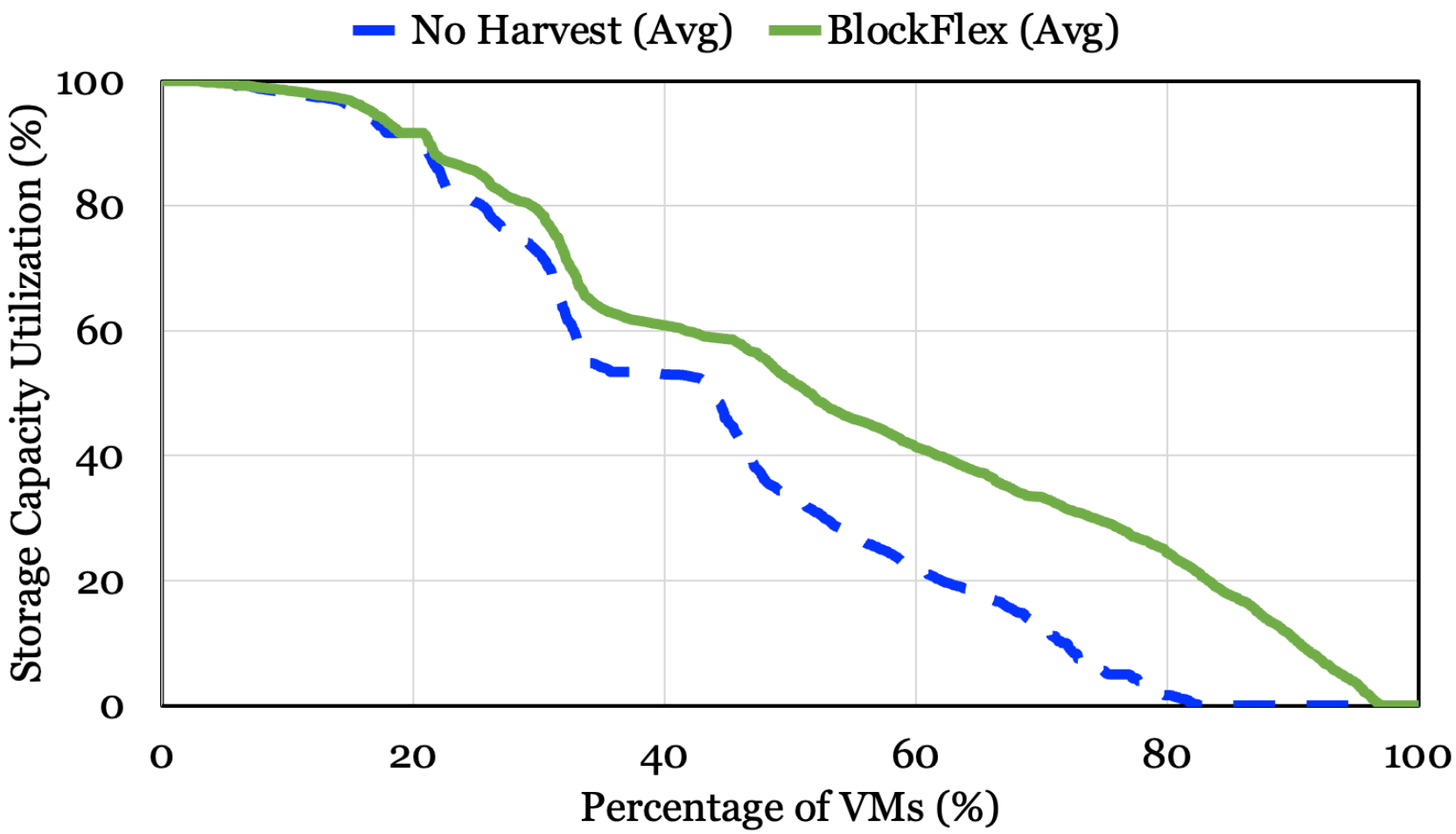
Regular VM Workloads

YCSB (Workloads A-E)

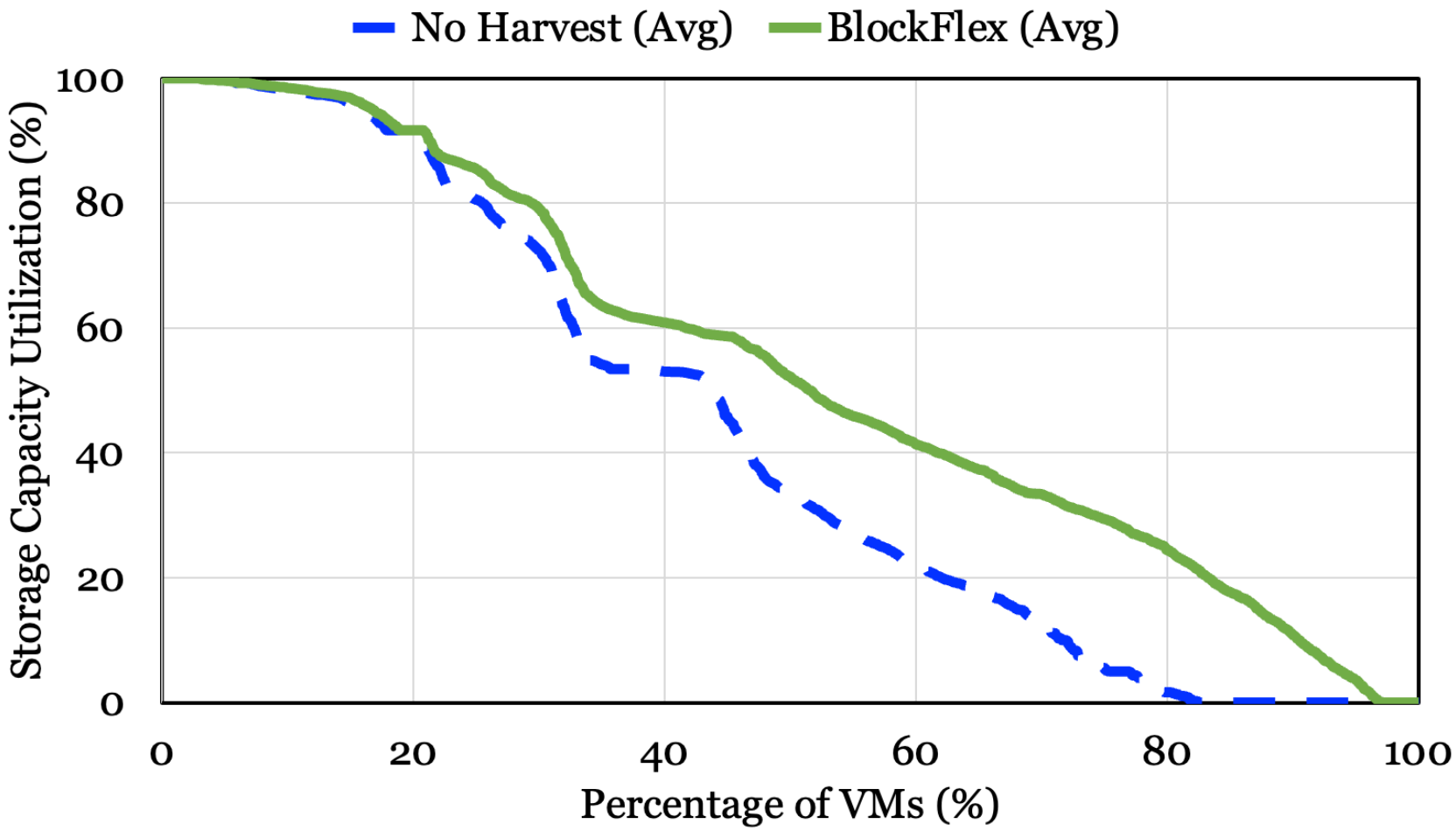
Harvest Workloads

Hadoop TersaSort
Graphchi PageRank
ML Image Preprocessing

Improved Cloud Storage Utilization with BlockFlex

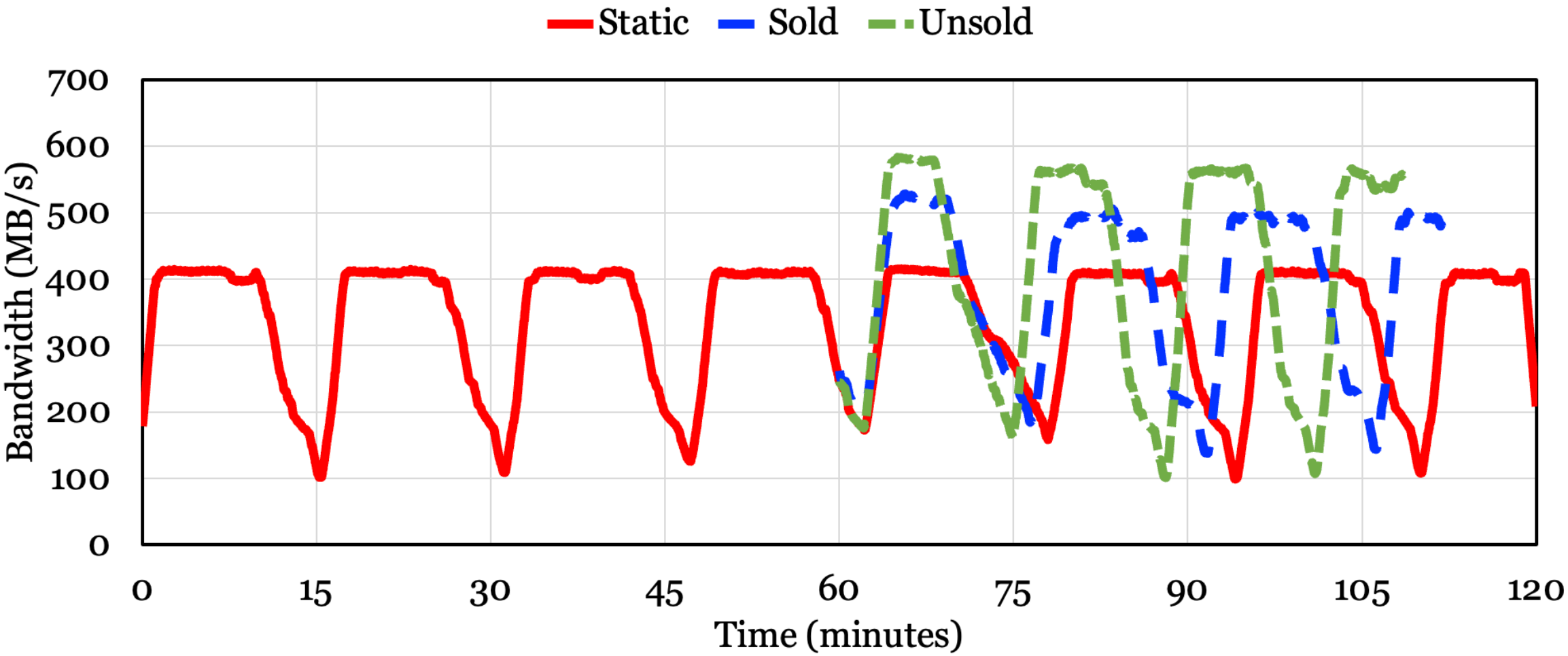


Improved Cloud Storage Utilization with BlockFlex

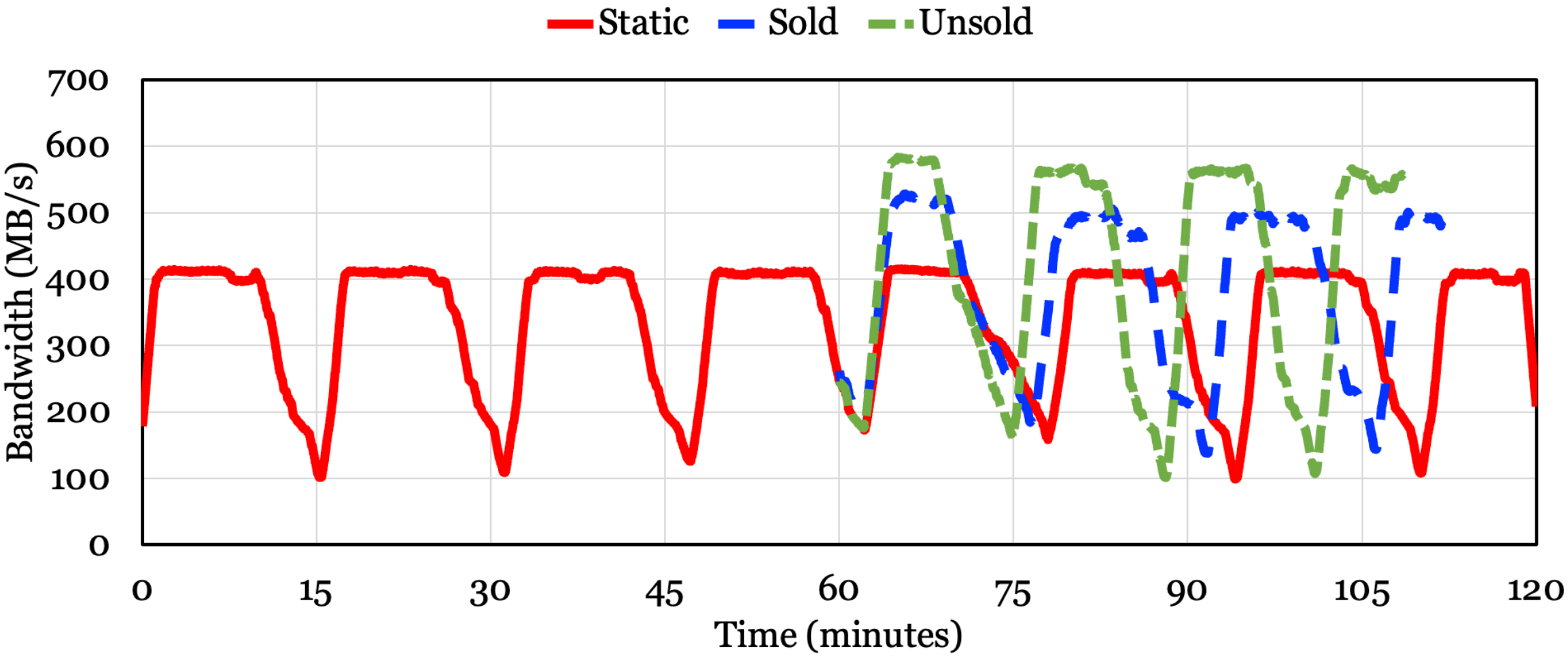


BlockFlex improves the storage utilization by 1.3x on average!

Improved Performance for Harvest VM with BlockFlex



Improved Performance for Harvest VM with BlockFlex



BlockFlex Improves Harvest VM Performance by up to 60%

BlockFlex Summary

Rethink Storage Virtualization for Storage Harvesting

BlockFlex Summary

Rethink Storage Virtualization for **Storage Harvesting**

Ensure Data Security

BlockFlex Summary

Rethink Storage Virtualization for **Storage Harvesting**

Ensure Data Security

Minimize the **Impact of Harvesting** on Regular VMs

BlockFlex Summary

Rethink Storage Virtualization for **Storage Harvesting**

Ensure Data Security

Minimize the **Impact of Harvesting** on Regular VMs

https://github.com/platformxlab/blockflex

Thank You!

Benjamin Reidys Jinghan Sun

Anirudh Badam Shadi Noghabi Jian Huang

breidys2@illinois.edu

Systems Platform Research Group

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN